



Graduate Theses, Dissertations, and Problem Reports

2017

Geographic Biases in Human Lyme disease Surveillance and Utility of Companion Animal Sentinel Programs: Exploratory Spatial Analysis

Brian Hendricks

Follow this and additional works at: <https://researchrepository.wvu.edu/etd>

Recommended Citation

Hendricks, Brian, "Geographic Biases in Human Lyme disease Surveillance and Utility of Companion Animal Sentinel Programs: Exploratory Spatial Analysis" (2017). *Graduate Theses, Dissertations, and Problem Reports*. 5793.

<https://researchrepository.wvu.edu/etd/5793>

This Dissertation is protected by copyright and/or related rights. It has been brought to you by the The Research Repository @ WVU with permission from the rights-holder(s). You are free to use this Dissertation in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you must obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/ or on the work itself. This Dissertation has been accepted for inclusion in WVU Graduate Theses, Dissertations, and Problem Reports collection by an authorized administrator of The Research Repository @ WVU. For more information, please contact researchrepository@mail.wvu.edu.

**Geographic Biases in Human Lyme disease Surveillance and Utility of Companion Animal
Sentinel Programs: Exploratory Spatial Analysis**

Brian Hendricks

Dissertation Submitted to the School of Public Health at West Virginia University
In partial fulfillment of the requirements for the degree of Doctor in Philosophy in Epidemiology

Thomas Hulsey, Ph.D., Committee Chair

Kim Innes, Ph.D., Faculty Advisor and Committee Co-Chair

Miguella Mark-Carew, Ph.D.

Jamison Conley, Ph.D.

Department of Epidemiology

Morgantown, West Virginia

2017

Keywords: Lyme disease, Spatial Analysis, Surveillance Systems, West Virginia

Copyright 2017 Brian Hendricks

ABSTRACT

Geographic Biases in Human Lyme disease Surveillance and Utility of Companion Animal Sentinel Programs: Exploratory Spatial Analysis

Brian Hendricks

Background and Objectives: Mapping and exploratory spatial data analyses are ideal tools for characterizing spread and occurrence of human Lyme disease infection. Unfortunately, many mapped displays utilizing Lyme disease surveillance data are prone to bias due to a lack of consideration for geographical confounders. The objectives of our study were to 1) characterize the geographic effects that boundary and travel related biases have on visualization of human Lyme disease occurrence and 2) apply these findings to develop a more precise methodology for evaluating efficacy of animal sentinel surveillance programs in predicting incidence of human Lyme disease infection.

Methods: County-level human Lyme disease and companion animal tick surveillance data were obtained from relevant state health departments. Data were organized within Microsoft Excel spreadsheets, and sorted by relevant reporting year and county. In Study 1, boundary effects were evaluated for the region containing Kentucky, Maryland, Ohio, Pennsylvania, Virginia, and West Virginia 2010-2014, utilizing a combination of rate smoothing and local indicators of spatial autocorrelation. Trends in disease clustering over time within our multistate region were evaluated utilizing logistic generalized estimating equations. In Study 2, travel associated biases were evaluated only for West Virginia confirmed Lyme disease cases 2011-2015, utilizing a combination of paired t-test, Wilcoxon Rank Signed test, and local indicators of spatial autocorrelation. In Study 3, the efficacy of the companion animal (dog and cat) sentinel surveillance program in West Virginia 2014-2016, was evaluated utilizing a combination of ordinary least squares and spatial regression techniques as well as local indicators of spatial autocorrelation on regression residuals.

Results: Study 1. Analyses indicated statistically significant ($P = 0.05$) clustering of human Lyme disease incidence over time. High-high clusters aggregated near counties bordering high incidence states, while low-low clusters aggregated near shared county borders in non-high incidence states. Study 2. Analyses indicated statistical non-equivalency using paired t-test ($t = 3.99$, $df = 54$, $P = 0.0002$) and the non-parametric Wilcoxon Signed Rank test ($S=264$, $P < 0.001$) between total overall cases and those obtained within patient's home county, suggesting significant travel-associated bias. Additionally, local indicators of spatial autocorrelation detected statistically significant ($P = 0.05$) patterns of clustering in the county level proportion of cases attributable to travel. Study 3. Regression analyses identified significant associations between confirmed cases of human Lyme disease and average number of *Ixodes scapularis* removed from dogs (ordinary least squares ($\beta=0.20$ $P < 0.001$) and spatial lag ($\beta = 0.12$, $P = 0.002$) models) but not cats for the period 2014-2016. Local indicators of spatial autocorrelation produced for spatial lag regression residuals indicated a decrease in model over and

underestimation, but identified a higher number of statistically significant outliers than ordinary least squares regression.

Conclusions: Results of spatial and regression analyses 1) indicate significant differential clustering of incident human Lyme disease within WV and surrounding states over time; 2) suggest substantial travel-associated bias in Lyme disease case visualization within WV; and 3) strongly support the use of companion animal, and specially dog sentinel surveillance programs for estimation of human Lyme disease risk within WV. These findings suggest that geographic biases significantly affect visualization of human Lyme disease incidence and support the effectiveness of utilizing dogs as sentinel populations to estimate human risk. Findings of these three studies highlight the importance of using statistical methodologies that can accommodate the spatial structure imbedded within public health surveillance data.

ACKNOWLEDGEMENTS

First and foremost, I thank West Virginia's School of Public Health, Department of Epidemiology and West Virginia Department of Health and Human Resources Bureau of Public Health for mentorship, funding, and resources. At this time, I would also like to thank Drs. Tom Hulse, Kim Innes, Miguella Mark-Carew, and Jamison Conley whose contributions and ideals led to completion of this project.

I am especially grateful for the support of my close friends Cara Mangine and Sara Warfield. Above all, I am most thankful for the love and support of my fiancé Kassey Riffle, parents Steve and Neana Hendricks, and grandparents Carl and Laura Hendricks who inspire me to continually set new personal and professional goals and seek out new adventures!

TABLE OF CONTENTS

Introduction	Pages	1	to	8
Chapter One	Pages	9	to	26
Chapter Two	Pages	27	to	34
Chapter Three	Pages	35	to	49
Project Conclusions	Pages	40	to	53

LIST OF TABLES

Table 1.1	Annual Confirmed and Probable Lyme disease Case Counts for Kentucky, Maryland, Ohio, Pennsylvania, Virginia, and West Virginia	Page 22
Table 3.1	Summary of Animal Species Specific Regression Models	Page 48

LIST OF FIGURES

Figure 1.1	Lyme disease Case Classification Criteria Flow Diagram	Page 21
Figure 1.2	Choropleth Maps Displaying Crude Yearly Incidence of Lyme Disease per 100,000 by County for Kentucky, Maryland, Ohio, Pennsylvania, Virginia, and West Virginia	Page 23
Figure 1.3	Choropleth Maps Displaying Spatial Empirical Bayesian Smoothing of Yearly Estimated Incidence per 100,000 by County for Kentucky, Maryland, Ohio, Pennsylvania, Virginia, and West Virginia	Page 24
Figure 1.4	Local Indicators of Spatial Autocorrelation (LISA) Cluster Analysis by county for Kentucky, Maryland, Ohio, Pennsylvania, Virginia, and West Virginia	Page 25
Figure 1.5	Proportion of County Level Lyme disease Reporting Hotspots (High-High), Cold spots (Low-Low), and Statistically Significant outliers (Low-High) within Kentucky, Maryland, Ohio, Pennsylvania, Virginia, and West Virginia	Page 26
Figure 2.1	Overall Total and Total Obtained within Patients Home County Confirmed Human Lyme disease Cases in West Virginia, 2011-2015	Page 33
Figure 2.2	Univariate Local Moran's I with Empirical Bayesian Rate. Local Indicators of Spatial Autocorrelation Maps Output From	Page 34
Figure 3.1	Univariate Local Moran's I for Dog and Cat Ordinary Least Squares and Spatial Lag Regression Residuals	Page 49

LIST OF NOMENCLATURE AND ABBREVIATIONS

Exploratory Spatial Data Analysis (ESDA)

Borrelia burgdorferi

Ixodes scapularis

Ixodes pacificus

Erythema Migrans (EM) Rash

Centers for Disease Control and Prevention (CDC)

Local Indicators of Spatial Autocorrelation (LISA)

Council of State and Territorial Epidemiologists (CSTE)

Odds Ratio (OR)

Generalized Estimating Equations (GEE)

INTRODUCTION

Lyme disease is the most common vectorborne disease in the United States (1, 2). Transmission of the etiological agent *Borrelia burgdorferi* to humans is accomplished through the prolonged attachment of the eastern blacklegged tick *Ixodes scapularis*, or western blacklegged tick *Ixodes pacificus* (3, 4). Clinical diagnosis of Lyme disease within endemic areas is primarily guided by the manifestation of the EM rash following a seven day incubation period (5). However, up to as many as 20-30% of patients infected with Lyme disease do not display the characteristic rash potentially resulting in misdiagnosis (6, 7). If left untreated, patients can develop severe chronic multisystem conditions with limited treatment options (5, 8).

Cases of human Lyme disease have been reported within every state of the contiguous United States (9), with two primary disease foci located in the Northeastern and North Central regions (10). Human case numbers vary dramatically by state, and often display a range of variable symptoms following infection (11, 12). As a result, criteria were established by Council of State and Territorial Epidemiologists to collect and analyze clinical, epidemiological, and laboratory diagnostic evidence within a national human Lyme disease surveillance system. Surveillance begins when potentially infected patients seek medical attention. Medical providers examine patients for physical manifestations of Lyme disease, and ask questions relating to exposure to tick habitat and recent travel history to ascertain probability of exposure (13). Physicians order diagnostic testing, which involves a screening test (ELISA or IFA), followed by a confirmatory test (IgG or IgM western blot) using the CDC approved two tier screening approach (13). This clinical, epidemiologic, and diagnostic evidence is then used to ascertain a case status (confirmed, probable, suspect, or not a case) to a patient (13).

Human exposure to Lyme disease is dependent on the overlapping distribution of ticks, hosts and favorable environmental factors described within the enzootic cycle of Lyme disease (14-16).

Ticks become infected as nymphs or larvae after ingesting blood from reservoir hosts, such as white footed mice (*Peromyscus leucopus*). In the traditional enzootic cycle, blacklegged ticks are infected as larvae and carry the bacteria onto each of their successive life stages (17). The majority of human cases of Lyme disease are reported in the spring and early summer months during peak nymphal activity (16). In the fall, adult blacklegged ticks seek large mammalian hosts such as white tailed deer (*Odocoileus virginianus*) to locate other ticks for mating and to consume their final bloodmeal. Humans and other dead end hosts such as dogs, cats, and horses can also become infected (17, 18) in the fall following attachment of adult blacklegged ticks (16).

Monitoring of enzootic transmission cycles for Lyme disease is inexact, as human exposure to infected ticks can vary depending on the climate, host availability, and tick specific diapause mechanisms (19). Uncertainty regarding the level of human exposure is further compounded by variability in Lyme disease reporting on a state by state basis (12). Public health researchers have attempted to pair human and animal Lyme disease surveillance data with exploratory spatial data analyses (ESDA) to identify high-risk groups and provide a clearer understanding of human risk. Common ESDA approaches used in public health include disease mapping, cluster analyses, and ecological analysis (20). Disease mapping applies statistical approaches to remove excess spatial variation from incidence maps (20-22). Ecological analyses investigate associations between the geographic distribution of disease incidence and potential explanatory factors (20, 23, 24). Disease clustering assesses the presence of abnormal aggregations of disease cases within mapped surveillance data (20, 21, 23). Together these approaches form the foundation of

exploratory spatial data analysis (ESDA) and provide a pathway to more complex modelling approaches (25).

Mapping is one of the most widely used types of graphical displays applied to surveillance data (23). However generalization of this technique among untrained users often results in inaccurate interpretation of mapped data (26, 27). Communicating map information in an interpretable and representative fashion is often complicated by cartographic principles such as scale, symbolization, and the need for further processing (27, 28). The map's scale dictates the extent to which data are averaged across the study area (28). Inappropriate choice of scale can mask or distort true geographic associations (22, 29). Likewise, in choosing semiotic display in maps, particular elements of the data are emphasized, while others are obscured or completely discarded (30). Lack of further processing is a common conceptual problem resulting from unavailability of data in the form most efficient for interpretation (28). Together these principles influence perception of mapped surveillance data. For example, incidence of disease is often displayed within thematic maps containing multiple predefined classes and color schemes. Audiences are left to their own conclusions regarding the maps overall message depending on the chosen color scheme and the range of values within the defined classes. These individualistic interpretations are potentially prone to biases without complementary statistical information (28).

Exploratory spatial analyses have been applied to investigate spatial trends in human and animal Lyme disease surveillance data. Findings focused on human surveillance data describe potential clustering and temporal trends in visualization of Lyme disease incidence (10, 31, 32), and varying model reliability between wild caught ticks and human infection by locality (33). While previous animal ESDA research explore associations between serosurveys and tick collections among dogs and human Lyme disease incidence (34-36). Findings indicated that boundary

effects and travel-associated infections potentially bias mapped representations of human incidence (20, 33, 37-39) and that companion animals such as dogs are useful indicators of high-risk for human populations (36, 40). However, studies that adjust for potential geographic biases in human studies, that incorporate analysis of time trends, and that adjust for spatial dependence or consider model efficacy in animal surveillance studies remain sparse. Consequently, the underlying spatial processes influencing the occurrence of human cases and the distribution of infected ticks and hosts remain obscured, resulting in potentially inaccurate representations of true human Lyme disease risk.

To accommodate spatial structure within public health Lyme disease surveillance data and fill current gaps in knowledge, we have developed a series of three complementary studies. The first study examines human Lyme disease reporting on a regional level and incorporates ESDA rate smoothing approaches to mitigate boundary effects. Controlling for boundary effects reduces the influence extreme differences in state level Lyme disease incidence have on accurate representation of county-level human Lyme disease risk. The second study evaluates travel-associated disease acquisition as a potentially important source of bias in visualization of human Lyme disease surveillance data in West Virginia and identifies spatial patterns in the proportion of cases attributable to travel among confirmed human Lyme disease cases. Our final investigation utilizes ordinary least squares and spatial regression models to identify significant associations between *Ixodes scapularis* companion animal collections and human Lyme disease incidence in West Virginia. We incorporate our research findings from the first two studies by removing travel-associated cases and adjusting for spatial dependence in county level human Lyme disease incidence. Additionally we evaluate model efficacy utilizing a combination of traditional and spatial model selection techniques. Together, these studies address important gaps

in current knowledge regarding the visualization of human and animal Lyme disease surveillance data.

LITERATURE CITED

1. CDC. How Many People Get Lyme Disease. 2015. (<http://www.cdc.gov/lyme/stats/humancases.html>). (Accessed 8/11/2016).
2. Hinckley A, Connally N, Meek J, et al. Lyme Disease Testing by Large Commercial Laboratories in the United States. *Clinical Infectious Diseases* 2014;59(5):676-81.
3. Parola P., Raoult D. Ticks and tickborne bacterial diseases in humans: an emerging infection threat. *Clin Infect Dis* 2001;32:897-928.
4. Estrada-Pena A. JF. Ticks feeding on humans: a review of records on human - biting Ixodidea with special reference to pathogen transmission. *Exp Appl Acarol* 1999;23:685-715.
5. Stanek G, Wormser G, Gray J, et al. Lyme borreliosis. *The Lancet* 2012;379:461-73.
6. Hayes E, Piesman J. How Can We Prevent Lyme Disease. *The New England Journal of Medicine* 2003;348:2424-30.
7. Bacon R, Kugeler K, Mead P. Surveillance for Lyme Disease --- United States, 1992--2006. *MMWR* 2008;57(SS10):1-9.
8. Li J, Hong K, Duan Y, et al. Spatial and Temporal Emergence Pattern of Lyme Disease in Virginia. *American Journal Tropical Medicine and Hygiene* 2014;91(6):1166-72.
9. CDC. Lyme disease data tables. 2015. (<http://www.cdc.gov/lyme/stats/tables.html>). (Accessed 8/11/2016).
10. Kugeler KJ., Farley GM., Forrester JD., et al. Geographic Distribution and Expansion of Human Lyme Disease, United States. *Emerging Infectious Diseases* 2015;21(8):1455-57.
11. Tibbles CD., Edlow JA. Does This PATient Have Erythema Migrans? *JAMA* 2007;297(23):2617-27.
12. Stromdahl E, Hickling G. Beyond Lyme: Etiology of Tick-borne Human Disease with Emphasis on the South-Eastern United States. *Zoonoses and Public Health* 2012;59(Suppl 2):48-64.
13. Centers for Disease Control and Prevention (CDC). Lyme Disease Case Ascertainment Criteria. 2017. (<https://wwwn.cdc.gov/nndss/conditions/lyme-disease/case-definition/2017/>). (Accessed April 2, 2017).
14. Ostfeld RS., Glass GE., Keesing F. Spatial epidemiology: an emerging (or re-emerging) discipline. *TRENDS in Ecology and Evolution* 2005;20(6):328-36.
15. Killilea M, Swei A, Lane R, et al. Spatial Dynamics of Lyme Disease: A Review. *EcoHealth* 2008;5:167-95.
16. Mullen GR., Durden LA. *Medical and Veterinary Entomology*. Burlington, MA: Academic Press Elsevier; 2009.
17. Radolf JD, Caimano, M.J., Stevenson, B., Hu, L.T.,. Of ticks, mice and men: understanding the dual-host lifestyle of Lyme disease spirochaetes [electronic article]. *Nat Rev Microbiol*.
18. Little SE., Heise SR., Blagburn BL., et al. Lyme Borreliosis in dogs and humans in the USA. *Trends in Parasitology* 2010;26(4):213-8.
19. Piesman J., Gern L. Lyme borreliosis in Europe and North America. *Parasitology* 2004;129:S191-S220.
20. Lawson AB. *Statistical Methods in Spatial Epidemiology*. Chichester, NY: Wiley & Sons Ltd.; 2013.

21. Razaeeian M., Dunn G., St Leger S., et al. Geographical epidemiology, spatial analysis, and geographical information systems: a multidisciplinary glossary. *Journal of Epidemiology and Community Health* 2007;61(2):98-102.
22. Elliot P., Wartenburg D. Spatial Epidemiology: Current Approaches and Future Challenges. *Environmental Health Perspectives* 2004;112(9):998-1006.
23. Lawson AB., Biggeri A., Bohning D., et al. *Disease Mapping and Risk Assessment For Public Health*. Chichester, NY: John Wiley & Sons Ltd; 1999.
24. Gatrell AC., Bailey TC. Interactive Spatial Data Analysis in Medical Geography. *Soc Sci Med* 1996;42(6):843-55.
25. Anselin L, Syabri I, Kho Y. GeoDa: An Introduction to Spatial Data Analysis. *Geospatial Analysis* 2006;38:5-22.
26. Jacquez G. Spatial analysis in epidemiology: Nascent science or a failure of GIS? *J Geography Sys* 2000;2:91-7.
27. Slocum TA., McMaster RB., Kessler FC., et al. *Thematic Cartography and Geovisualization*. Upper Saddle River, NJ: Pearson Prentice Hall; 2009.
28. Lawson AB. *Statistical Methods in Spatial Epidemiology*. West Sussex, England: John Wiley & Sons Ltd.; 2001.
29. Lomolino MV., Riddle BR., Whittaker RJ., et al. *Biogeography*. Sunderland, MA: Sinauer Associates Inc.; 2010.
30. Carpendale MST. *Considering Visual Variables as a basis for Information Visualization*. University of Calgary: Department of Computer Science; 2003.
31. Guo L., Sun Z., Di L., et al. Spatial Distribution and variation analysis of Lyme disease in the Northeastern United States. *Agro-Geoinformatics* 2016;Fifth International Conference on IEEE:1-4.
32. Diuk-Wasser M, Gatewood A, Cislo P, et al. Human Risk of Infection with *Borrelia burgdorferi*, the Lyme Disease Agent, in Eastern United States. *American Journal of Tropical Medicine and Hygiene* 2012;86(2):320-7.
33. Pepin KM., Eisen RJ., Mead PS., et al. Geographic Variation in the Relationship between Human Lyme Disease Incidence and Density of Infected Host Seeking *Ixodes scapularis* Nymphs in the Eastern United States. *American Journal of Tropical Medicine and Hygiene* 2012;86(6):1062-71.
34. Glickman LT., Moore GE., Glickman NW., et al. Purdue University-Banfield National COmpanion Animal Surveillance Program for Emerging and Zoonotic Diseases. *Vector-Borne and Zoonotic Diseases* 2006;6(1):14-23.
35. Duncan Aw., Correa MT., Levine JF., et al. The Dog as a Sentinel for Human Infection: Prevalence of *Borrelia burgdorferi* C6 Antibodies in Dogs from Southeastern and Mid-Atlantic States. *Vector-Borne and Zoonotic Diseases* 2005;5(2):101-9.
36. Smith FD., Ballantyne R., Morgan ER., et al. Estimating Lyme disease risk using pet dogs as sentinels. *Comparative Immunology, Micorbiology, and Infectious Diseases* 2012;35:163-7.
37. Steere AC., Malawista SE. Cases of Lyme Disease in the United States: Location Correlated with Distribution of *Ixodes dammini*. *Annals of Internal Medicine* 1979;91:730-3.
38. Hendricks B., Mark-Carew M. using exploratory data analysis to identify and predict patterns of human Lyme disease case clustering within a multistate region, 2010-2014. *Spatial and Spatio-temporal Epidemiology* 2017;20:35-43.

39. Forrester J, Brett M, Matthias J, et al. Epidemiology of Lyme disease in low-incidence states. *Ticks and Tick-borne Diseases* 2015;6:721-23.
40. Hamer SA., Tsao JI., Walker ED., et al. Use of tick surveys and serosurveys to evaluate pet dogs as a sentinel species for emerging Lyme disease. *AJVR* 2009;70(1):49-56.

CHAPTER ONE

Using Exploratory Data Analysis to Identify and Predict Patterns of Human Lyme Disease Case Clustering within a Multistate Region, 2010-2014

ABSTRACT

Lyme disease is the most commonly reported vectorborne disease in the United States. The objective of our study was to identify patterns of Lyme disease reporting after multistate inclusion to mitigate potential border effects. County-level human Lyme disease surveillance data were obtained from Kentucky, Maryland, Ohio, Pennsylvania, Virginia, and West Virginia state health departments. Rate smoothing and Local Moran's I was performed to identify clusters of reporting activity and identify spatial outliers. A logistic generalized estimating equation was performed to identify significant associations in disease clustering over time. Resulting analyses identified statistically significant ($P=0.05$) clusters of high reporting activity and trends over time. High reporting activity aggregated near border counties in high incidence states, while low reporting aggregated near shared county borders in non-high incidence states. Findings highlight the need for exploratory surveillance approaches to describe the extent to which state level reporting affects accurate estimation of Lyme disease progression.

Keywords: Lyme disease; Local Moran's I ; Surveillance

INTRODUCTION

Lyme disease is the most common vectorborne disease in the United States with an estimated 300,000 new individuals effected annually (1, 2). Acute human infection is typically associated with one or more erythema migrans (EM) (bull's eye rash), arthritis, and to a lesser extent Lyme carditis or neuroborreliosis (3). While early antibiotic therapy is effective in up to 90% of diagnosed patients (4), approximately 20-30% of patients do not exhibit an EM rash complicating early diagnosis (5, 6). As a result, many possible human Lyme infections are misdiagnosed or left untreated, potentially resulting in chronic debilitating multisystem conditions with limited treatment options (3, 7).

Human Lyme disease infection has been reported from every state in the United States except Hawaii (8). State level case counts include both locally acquired and travel associated. In recent years there has been an increasing trend in Lyme disease case reporting; case counts vary by state, and are thought to be associated with increasing abundance and changing distributions of tick vectors (9). This increase in the number of human infections paired with changing native tick populations emphasize the need for sustainable surveillance strategies focused on targeted prevention and early disease recognition and treatment programs (10).

Exploratory spatial data analysis (ESDA) is one such strategy, allowing analysts the ability to explore data in terms of space and time. Common ESDA techniques used in public health include raw and spatial rate smoothing as well as cluster analysis (11-14). Spatial rate smoothing can be used to increase interpretability of disease maps by adjusting for variations in rate estimation (7). Cluster analyses are employed to investigate whether disease occurred by random chance, and understand directionality of disease expansion (7).

These techniques are ideal Lyme disease surveillance tools because of the inherent spatial processes that drive human pathogen transmission (15). Infected ticks must be in close spatial proximity to humans for transmission to occur. Previous studies have employed ESDA to visualize rates of human Lyme disease and identify pockets of high disease intensity. The authors found that cases of Lyme disease appeared to cluster near or at state lines, and believed this to be the result of misdiagnosis, disproportionate presence of infected ticks, or travel associated infections (7, 13, 14, 16, 17). These possible explanations are supported by previous work which describe the distribution of Lyme disease between states as dynamic, and specifically identify travel as potential source for ~80% of cases reported within low incidence states (9, 18).

The region of the United States that includes the Midwest, Mid-Atlantic, and Southeastern states of Kentucky, Maryland, Ohio, Pennsylvania, Virginia, and West Virginia is an ideal study area to characterize Lyme disease case reporting near or at state lines. High incidence states within the region include Maryland, Pennsylvania, and Virginia, all of which are in the top 15 states that report Lyme in the nation (19). Alternatively, Kentucky, Ohio, and West Virginia are in the mid to low range for Lyme disease incidence, but are bordered by one or more of the high incidence states within our study region (19). To date no known study has described the impact multistate inclusion has on geographic visualization of Lyme disease incidence within a region, as opposed to a state by state basis. The objective of our study was to identify patterns of Lyme disease reporting after multistate inclusion to mitigate potential border effects.

METHODS

Data Acquisition and Management

County-level probable and confirmed human Lyme disease surveillance data for 2010-2014 were obtained from the state health departments in West Virginia, Virginia, Ohio, Pennsylvania, Kentucky, and Maryland. Complete annual case totals were received from all states except for Kentucky, which had reporting data from 2012-2014. Surveillance data used in our analysis were limited to confirmed and probable case counts as those are what are nationally reported. Data were specifically requested from 2010-2014 to incorporate stricter diagnostic criteria within 2008 and 2011 Lyme case definitions (7). Utilizing stricter criteria for case ascertainment led to increased consistency of medical, laboratory, and epidemiologic evidence among reported cases. Lyme disease case classification standards at each of the respective state health departments within the region were based on the current Council of State and Territorial Epidemiologist Case Definition for Lyme disease (20). A decision diagram is presented in Figure 1.1 to summarize current Lyme disease case classification criteria (21).

Data were sorted in Microsoft Excel 2010 by state, county, and reporting year. Confirmed and probable Lyme disease case reports were paired with U.S. census county level population estimates by year and county (22). Excel spreadsheets containing the county level case information and population census data were joined within ArcGIS 10.1 (ESRI, Redlands CA) to a United States Counties and Localities shapefile obtained from U.S. Department of Agriculture National Resources Conservation Service Geospatial Gateway (USDA: NRCS) (23). The newly joined features were exported as a shapefiles and analyzed in GeoDa 1.6 (GeoDa Center, Tempe, AZ).

Descriptive analysis

Crude yearly estimated incidence rates were initially visualized to display variation in raw data. Estimated incidence by county was calculated by dividing the total number of confirmed and

probable Lyme disease cases by the respective county's yearly July 1st census estimates. Spatial weight connectivity histograms were constructed to characterize the distribution of neighbor structures for rook and queen spatial weights (24). Rook spatial weights characterize polygons which share a common border as neighbors, while queen spatial weights characterize all polygons which share a common border or vertex as neighbors. Typically, this difference results in a stricter neighbor criterion when rook spatial weights are applied. However, distribution of neighbors among contiguity weights in our study were almost identical, as a result the rook spatial weight was chosen due to the inconsistent size of counties within the region (13, 25). Smoothing techniques were applied to adjust crude incidence rates and reduce potential spurious outliers and biases (16). Adjusted yearly estimated incidence were calculated from the raw data after applying our rook spatial weight using Spatial Empirical Bayesian techniques to account for maximum localized instability of Lyme disease reporting rates among neighboring counties (26, 27). Raw and adjusted yearly estimated incidence rates were expressed in terms of cases per 100,000 persons using a quantile classification system with five groupings (0, 4, 20, 60, and >60 cases per 100,000 persons) within ArcGIS 10.1. Incidence rates were classified this way to place specific emphasis on areas with low to moderate occurrence of Lyme disease.

Cluster Analysis

Local Empirical Bayesian Moran's I estimates incorporating our previously defined rook spatial weight were conducted for each year independently by setting yearly annual total case count and yearly population census estimates as event and base variables respectively. Resulting Moran's I scatterplots and local indicators of spatial autocorrelation (LISA) cluster maps with Empirical Bayesian rates were permuted 99,999 times at a pseudo p-value of 0.05 to optimize sensitivity of clustering and detection of statistically significant outliers while accounting for instability of

population at risk for each year within the region (26, 28, 29). A generalized estimating equation (GEE) were calculated in SAS 9.4 (SAS Institute Inc., Cary NC) with a logit link, binomial distribution, and repeated measures statement to describe any significant relationships between clusters of high reporting activity and time. Our GEE model excluded counties within Ohio and Kentucky because there were no high-high clusters identified for any of the years within our study period.

RESULTS

Descriptive analysis

Overall from 2010-2014, 41,922 confirmed and probable Lyme disease cases were reported with 0.25% (n=103) in Kentucky, 17.1% (n=7,186) in Maryland, 0.90% (n=375) in Ohio, 66% (n=27,591) in Pennsylvania, 14.4% (n=6,031) in Virginia, and 1.52% (n=636) in West Virginia (Table 1.1). Yearly estimated incidence for states included within the region ranged from 17.1-23.3 per 100,000 persons. Among non-high incidence states, West Virginia had the highest estimated incidence of Lyme disease (5.22-7.34 per 100,000 persons), and accounted for only 4.14-4.64% of the regions estimated population. Alternatively, Ohio accounted for 25.8-28.9% of the regions estimated population, but only had an estimated incidence of 0.38-1.03 per 100,000 persons. Among the region's high incidence states, Pennsylvania had the highest estimated incidence of Lyme disease (29.9-58.9 per 100,000 persons) and accounted for the highest proportion of the regions estimated population (28.5-35.1%). While Maryland accounted for a lower proportion of individuals in high incidence states than Virginia (13.0-14.5% versus 18.4-20.2%), it had a higher estimated incidence 20.1-28.1 vs 12.6-16.2 per 100,000 persons respectively.

County-level Lyme disease incidence per 100,000 persons for each state included within our study region is displayed in Figure 1.2. Variation in county-level incidence was reduced following Spatial Empirical Bayesian (SEB) rate, particularly in counties where the population at risk was smaller (Figure 1.3). Yearly proportion of counties within the region with no reported cases of Lyme disease decreased from 45-55% to 8-24%. Alternatively, a smaller effect was seen in counties which reported a yearly incidence of > 60 cases per 100,000 persons (11.7-15.2% vs 11.7-13.8%). Despite this decrease, West Virginia counties of Berkeley, Hampshire, Jefferson, and Morgan remained the only counties to report an incidence of > 60 cases per 100,000 persons within a non-high incidence state in our study region.

Cluster analysis

Local Empirical Bayesian Moran's I estimates of spatial autocorrelation for 2010-2014 were significant ($P=0.05$) for each year, and ranged from 0.18 to 0.30. High-high clusters were identified in Maryland, Pennsylvania, Virginia, and the eastern panhandle of West Virginia (Figure 1.4). Alternatively, low-low clusters of Lyme reporting were identified in Kentucky, Ohio, south western Virginia, and West Virginia (Figure 1.4). Spatial outliers (low-high) ($n=30$) were identified in Maryland, Pennsylvania, Virginia, and West Virginia during the study period. Overall, the number of counties regarded as hot spots (high-high) increased over time, while cold spots (low-low), and statically significant spatial outliers (low-high) fluctuated depending on year (Figure 1.5). Generalized estimating equations (GEE) were used to assess the trend of increasing high-high clusters over time from 2010-2014. The fitted model shows that there is a significant trend (increasing frequency of clusters over time) with $OR = 1.29$ ($P < 0.001$). An exchangeable correlation structure was used in the GEE logistic model.

DISCUSSION

While spatial analyses have been used for some time in relevant literature, this is one of few studies to apply spatial analysis on a multistate scale to adjust for reporting activity in neighboring states. Additionally, this is the first known study to characterize county-level Lyme case clustering within Kentucky, Ohio, and West Virginia. Among states included, Kentucky had the lowest number of Lyme disease cases reported (14 to 45), while Pennsylvania had the highest (3,806 to 7,487). The greatest proportion of the multistate region's population resided in Pennsylvania (28.5 to 31.8%), while the smallest proportion resided in West Virginia (4.14 to 4.64%). Incidence rate visualization per 100,000 persons (Figures 1.2-1.3) revealed high reporting activity in the east and low reporting in the west. Both raw and smoothed rate visualization highlight the decreasing trend over time for counties with no reported cases of Lyme disease in the western part of the multistate region, excluding the increase within West Virginia and Ohio 2011-2012. However, smoothed incidence maps display reduced variability in county-level reporting among neighboring counties, and displayed fewer counties with no Lyme disease surveillance data. LISA cluster maps (Figure 1.4) indicated an increasing frequency of high-high counties observed over time in the eastern part of the multistate region with particular emphasis in the southwestern corner of Virginia and northeastern panhandle of West Virginia. This increase over time in high reporting activity within the multistate region is visualized in Figure 1.5 as the proportion of counties having high-high report clustering over time. Graphical representation of clustering within the region was displayed as a proportion of the total available county data due to lack of county level Lyme disease case data for Kentucky, 2010-2011. Subsequent GEE logistic analysis further supported our findings indicating increasing odds of county level high-high clustering over our study period within those states which had at least one high-high cluster identified in the LISA cluster map.

Study findings are based on a multistate exploratory spatial analysis of county level human Lyme disease case data. It is important to consider that while ESDA is particularly suited for identifying patterns of disease, it does not have the ability to explain the observed patterns (13). Other potential limitations associated with the use of surveillance data are linked to differential reporting biases which may potentially result in under or over reporting (30). Possible sources of differential reporting bias are varying presentations of Lyme disease signs and symptoms among infected patients, varying interpretations of the Lyme disease case definitions and unequal Lyme disease awareness among physicians and public health professionals (5, 6, 31). We attempt to adjust for varying interpretations of Lyme case definitions by limiting data acquisition to 2010-2014, a period in which Lyme case definitions were most similar (20). Unfortunately, there was no way to control for varying levels of Lyme disease awareness or physician reported Lyme patient symptoms because of the aggregate and retrospective nature of our investigation.

Despite these potential limitations, public health surveillance data provide valuable information in regards to Lyme disease occurrence, particularly in areas where pathogen prevalence among naturally occurring tick populations is rare or largely uncertain (32, 33). National Lyme disease surveillance data suggest a positive temporal association with cases of disease increasing over time with two prominent foci in the northeastern and north central United States (33, 34).

Counties within these high incidence foci have continually expanded to now include localities within Maryland, Pennsylvania, Virginia, and West Virginia (33). This nationally observed trend is consistent with our study findings as well as studies conducted at the regional and state level, noting high reporting activity in Pennsylvania and Maryland as well as increasing frequency of high reporting clusters in Virginia (7, 35).

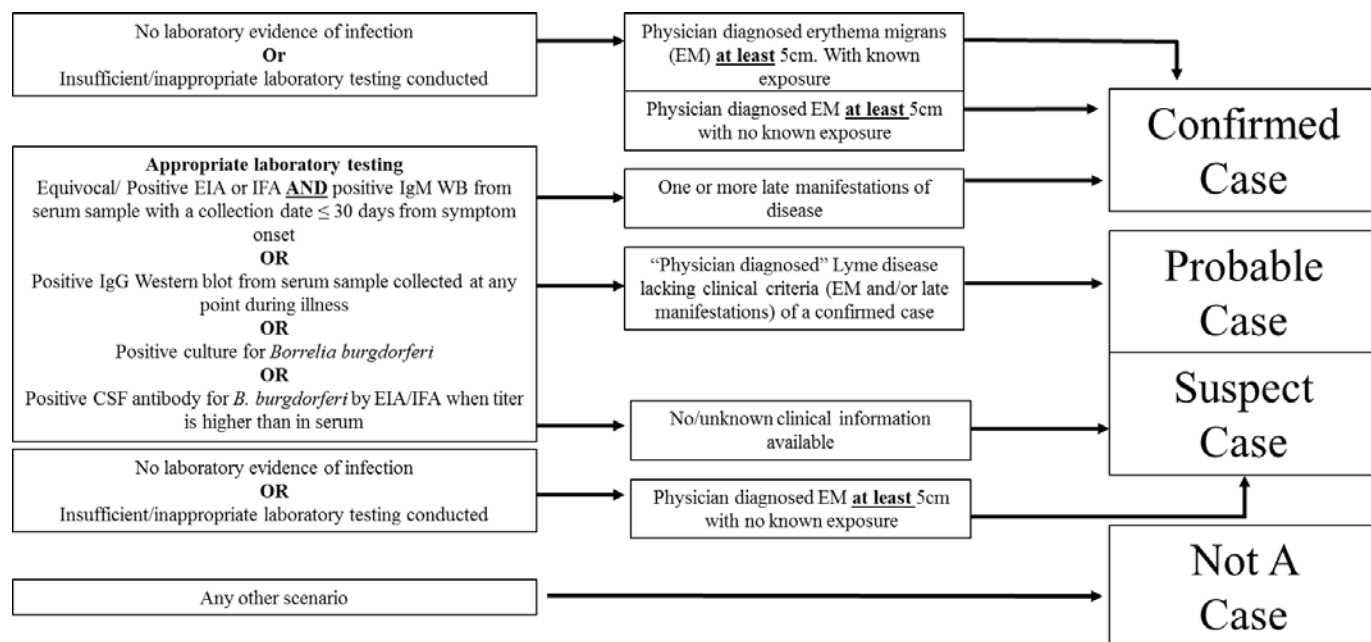
In general, epidemiologic patterns of human Lyme disease are largely dependent on tick-host environmental factors as well as human behaviors such as participation in outdoor activities, pet ownership, and travel (15, 36, 37). No human behavioral data were collected for any of the cases analyzed. However, spatial analysis did reveal clusters of high reporting activity near state lines in Maryland, Pennsylvania, Virginia, and West Virginia. This could suggest that travel history is important, particularly in non-high incidence states such as West Virginia, where Lyme ecology is largely undescribed. Unfortunately, there are a limited number of epidemiologic studies specifically focusing on Lyme disease occurrence in non-high incidence states. Exploratory spatial data analyses are particularly suited for public health professionals in these areas providing baseline risk assessment, and subsequent surveillance initiatives to illustrate spatio-temporal changes in disease occurrence. Further research into the utility of spatial analysis for public health data is warranted, particularly in non-high incidence states where Lyme disease occurrence is largely undescribed.

LITERATURE CITED

1. CDC. How Many People Get Lyme Disease. 2015. (<http://www.cdc.gov/lyme/stats/humancases.html>). (Accessed 8/11/2016).
2. Hinckley A, Connally N, Meek J, et al. Lyme Disease Testing by Large Commercial Laboratories in the United States. *Clinical Infectious Diseases* 2014;59(5):676-81.
3. Stanek G, Wormser G, Gray J, et al. Lyme borreliosis. *The Lancet* 2012;379:461-73.
4. Wormser G, Nadelman R, Dattwyler R, et al. Practice Guidelines for the Treatment of Lyme Disease. *Clinical Infectious Diseases* 2000;31(Suppl 1):S1-14.
5. Hayes E, Piesman J. How Can We Prevent Lyme Disease. *The New England Journal of Medicine* 2003;348:2424-30.
6. Bacon R, Kugeler K, Mead P. Surveillance for Lyme Disease --- United States, 1992--2006. *MMWR* 2008;57(SS10):1-9.
7. Li J, Hong K, Duan Y, et al. Spatial and Temporal Emergence Pattern of Lyme Disease in Virginia. *American Journal Tropical Medicine and Hygiene* 2014;91(6):1166-72.
8. CDC. Lyme disease data tables. 2015. (<http://www.cdc.gov/lyme/stats/tables.html>). (Accessed 8/11/2016).
9. Stromdahl E, Hickling G. Beyond Lyme: Etiology of Tick-borne Human Disease with Emphasis on the South-Eastern United States. *Zoonoses and Public Health* 2012;59(Suppl 2):48-64.
10. Atkinson S, Sarkar S, Avina A, et al. A determination of the spatial concordance between Lyme disease incidence and habitat probability of its primary vector *Ixodes scapularis* (black-legged tick). *Geospatial Health* 2014;9(1):203-12.
11. Waller L, Gotway C. *Applied Spatial Statistics for Public Health*. 1 ed. Hoboken, NJ: Wiley-Interscience; 2004.
12. Waller LA., Goodwin BJ., Wilson ML., et al. Spatio-temporal patterns in county-level incidence and reporting of Lyme disease in the northeastern United States, 1990–2000. *Environ Ecol Stat* 2007;18:83-100.
13. Szonyi B, Srinath I, Esteve-Gassent M, et al. Exploratory spatial analysis of Lyme disease in Texas -what can we learn from the reported cases? *BMC Public Health* 2015;15(924):1-8.
14. Brinkerhoff J, Gilliam W, Daines D. Lyme Disease, Virginia, USA, 2000-2011. *Emerging Infectious Diseases* 2015;20(10):1661-68.
15. Killilea M, Swei A, Lane R, et al. Spatial Dynamics of Lyme Disease: A Review. *EcoHealth* 2008;5:167-95.
16. Glavanakov S, White D, Caraco T, et al. Lyme Disease in New York State: Spatial Pattern at a Regional Scale. *American Journal of Tropical Medicine and Hygiene* 2001;65(5):538-45.
17. Xue L, Scoglio C, McVey S, et al. Two Introductions of Lyme Disease into Connecticut: A Geospatial Analysis of Human Cases from 1984 to 2012. 2015:523-28.
18. Forrester J, Brett M, Matthias J, et al. Epidemiology of Lyme disease in low-incidence states. *Ticks and Tick-borne Diseases* 2015;6:721-23.
19. CDC. Lyme disease data and statistics. 2015. (<https://www.cdc.gov/lyme/stats/index.html>). (Accessed 2/11/2016).
20. CDC. Lyme Case Definitions 2008 and 2011. 2011. (<https://wwwn.cdc.gov/nndss/conditions/lyme-disease/case-definition/2011/>). (Accessed 1/20/2016).

21. (WVBPH) WVBoPH. Lyme disease Case Ascertainment Guide. In: Carew MM, ed, 2016:Guide based on 2011 CSTE Lyme disease case classification criteria.
22. United States Census Bureau. Population Estimates. 2015. (<http://www.census.gov/popest/>). (Accessed 1/20/2016).
23. USDA NRCS. Geospatial Gateway. 2015. (<https://gdg.sc.egov.usda.gov/>). (Accessed 1/20/2016).
24. Anselin L, Syabri I, Kho Y. GeoDa: An Introduction to Spatial Data Analysis. *Geospatial Analysis* 2006;38:5-22.
25. GeoDa. Which Spatial Weight Should I Choose? ; 2016. (<https://geodacenter.github.io/glossary.html#esda>). (Accessed 9/10/2016).
26. Poh-Chin L, Fun-Mun S, Chan K. *Spatial Epidemiological Approaches in Disease Mapping and Analysis*. CRC Press; 2008.
27. Anselin L. GeoDa 0.9 user's Guide. In: Illinois Uo, ed. *Center for Spatially Integrated Social Science*, 2003.
28. Anselin L. Exploring Spatial Data with GeoDa: A Workbook. *University of Illinois Urbana-Champaign* 2005;Center for Spatially Intergated Social Science.
29. Assuncao RM., Reis EA. A New Proposal to Adjust Moran's I for Population Density. *Statistics in Medicine* 1999;18:2147-62.
30. Diuk-Wasser M, Gatewood A, Cislo P, et al. Human Risk of Infection with *Borrelia burgdorferi*, the Lyme Disease Agent, in Eastern United States. *American Journal of Tropical Medicine and Hygiene* 2012;86(2):320-7.
31. Brett M, Hinckley A, Zielinski-Gutierrez E, et al. U.S. healthcare providers' experience with Lyme and other tick-borne diseases. *Ticks and Tick-borne Diseases* 2014;5:404-8.
32. Pepin KM., Eisen RJ., Mead PS., et al. Geographic Variation in the Relationship between Human Lyme Disease Incidence and Density of Infected Host Seeking *Ixodes scapularis* Nymphs in the Eastern United States. *American Journal of Tropical Medicine and Hygiene* 2012;86(6):1062-71.
33. Kugeler KJ., Farley GM., Forrester JD., et al. Geographic Distribution and Expansion of Human Lyme Disease, United States. *Emerging Infectious Diseases* 2015;21(8):1455-57.
34. Eisen RJ., Eisen L., CB. B. County-Scale Distribution of *Ixodes scapularis* and *Ixodes pacificus* (Acari:Ixodidae) in the Continental United States. *Journal of Medical Entomology* 2016;0(0):1-38.
35. Guo L., Sun Z., Di L., et al. Spatial Distribution and variation analysis of Lyme disease in the Northeastern United States. *Agro-Geoinformatics* 2016;Fifth International Conference on IEEE:1-4.
36. Bunnell J, Price S, Das A, et al. Geographic Information Systems and Spatial Analysis of Adult *Ixodes scapularis* (Acari: Ixodidae) in the Middle Atlantic Region of the U.S.A. *Journal of Medical Entomology* 2003;40(4):570-76.
37. Eisen R, Lane R, Fritz C, et al. Spatial Patterns of Lyme Disease Risk in California Based On Disease Incidence Data and Modeling of Vector-Tick Exposure. *American Journal of Tropical Medicine and Hygiene* 2006;75(4):669-76.

Figure 1.1 Graphical Representation of Case Classification Criteria from the 2011 Council of State and Territorial Epidemiologists (CSTE) Lyme disease Case Definition.



** Graphic was obtained from the 2016 West Virginia Bureau of Public Health Lyme disease Case Ascertainment Toolkit.

Table 1.1 Annual Human Lyme Disease Case Counts within our Multi State Region of Kentucky, Maryland, Ohio, Pennsylvania, Virginia, and West Virginia, 2010-2014.

State	Population (%)	Case Investigations (%)	Yearly Estimated Incidence (per 100,000)
2010			
Region	39,922,622	6,836	17.1
Kentucky	-	-	-
Maryland	5,787,193 (14.5%)	1,617 (23.7%)	27.9
Ohio	11,545,435 (28.9%)	44 (0.64%)	0.38
Pennsylvania	12,710,472 (31.8%)	3,805 (55.7%)	29.9
Virginia	8,025,376 (20.1)	1,245 (18.2%)	15.5
West Virginia	1,854,146 (4.64%)	125 (1.82%)	6.74
2011			
Region	40,096,695	7,924	19.8
Kentucky	-	-	-
Maryland	5,840,241 (14.6%)	1,351 (17.0%)	23.1
Ohio	11,549,772 (28.8%)	53 (0.67%)	0.46
Pennsylvania	12,741,310 (31.8%)	5,362 (67.7%)	42.1
Virginia	8,110,188 (20.2%)	1,023 (12.9%)	12.6
West Virginia	1,855,184 (4.63%)	135 (1.70%)	7.27
2012			
Region	44,632,206	7,971	17.9
Kentucky	4,379,730 (9.81%)	14 (0.17%)	0.32
Maryland	5,884,868 (13.2%)	1,651 (20.7%)	28.1
Ohio	11,553,031 (25.9%)	66 (0.83%)	0.57
Pennsylvania	12,764,475 (28.6%)	5,033 (63.1%)	39.4
Virginia	8,193,422 (18.4%)	1,110 (13.9%)	13.5
West Virginia	1,856,680 (4.16%)	97 (1.22%)	5.22
2013			
Region	44,793,367	8,685	19.4
Kentucky	4,395,295 (9.81%)	44 (0.51%)	1.00
Maryland	5,928,814 (13.2%)	1,194 (13.7%)	20.1
Ohio	11,570,808 (25.8%)	93 (1.07%)	0.80
Pennsylvania	12,773,801 (28.5%)	5,904 (68.0%)	46.2
Virginia	8,270,345 (18.5)	1,307 (15.0%)	15.8
West Virginia	1,854,304 (4.14%)	143 (1.65%)	7.71
2014			
Region	44,531,085	10,387	23.3
Kentucky	4,339,367 (9.74%)	45 (0.43%)	1.03
Maryland	5,773,552 (13.0%)	1,373 (13.2%)	23.8
Ohio	11,536,504 (25.9%)	119 (1.14%)	1.03
Pennsylvania	12,702,379 (28.5%)	7,487 (72.1%)	58.9
Virginia	8,326,289 (18.7%)	1,346 (13.0%)	16.2
West Virginia	1,852,994 (4.16%)	136 (1.31%)	7.34

** Population data are July 1st mid-year population estimates from each year obtained from the U.S Census Bureau at the county level.

** Annual Case counts are the sum of the confirmed and probable cases reported by each state's respective state level health department. No case investigation data could be gathered from the Kentucky State Department of Health for 2010 and 2011. As a result population data for Kentucky are also excluded to make our descriptive analysis as representative of available data within the region as possible.

Figure 1.2 Quantile Choropleth Maps Displaying Yearly Crude Incidence by County per 100,000 Persons for Kentucky, Maryland, Ohio, Pennsylvania, Virginia, and West Virginia. No data are present for Kentucky from 2010-2011, due to data quality issues self-reported by the Kentucky State Department of Health during data acquisition.

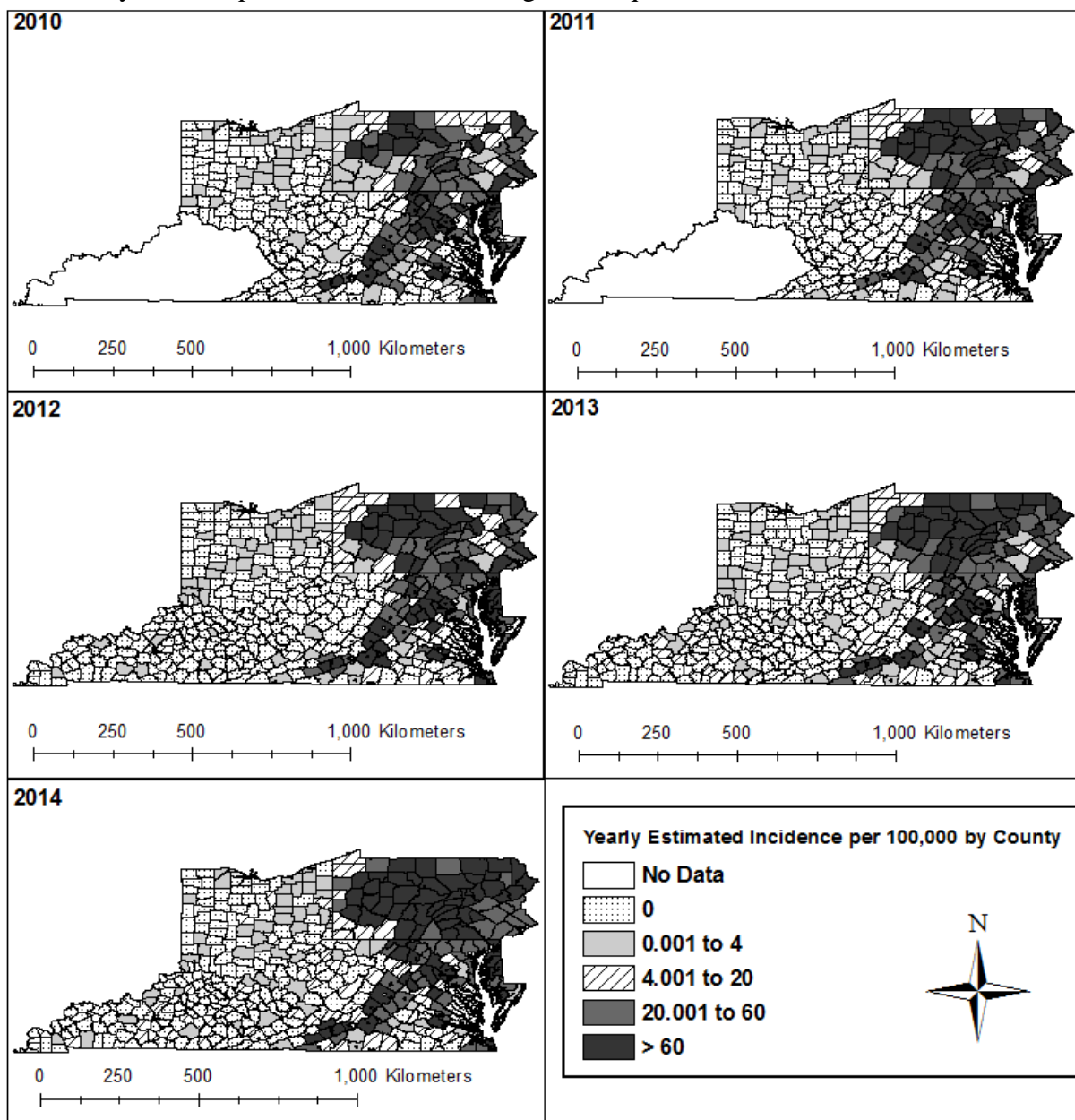


Figure 1.3 Spatial Empirical Bayesian Smoothed Quantile Choropleth Maps Displaying Yearly Adjusted Incidence by County per 100,000 Persons for Kentucky, Maryland, Ohio, Pennsylvania, Virginia, and West Virginia. No data are present for Kentucky from 2010-2011, due to data quality issues self-reported by the Kentucky State Department of Health during data acquisition.

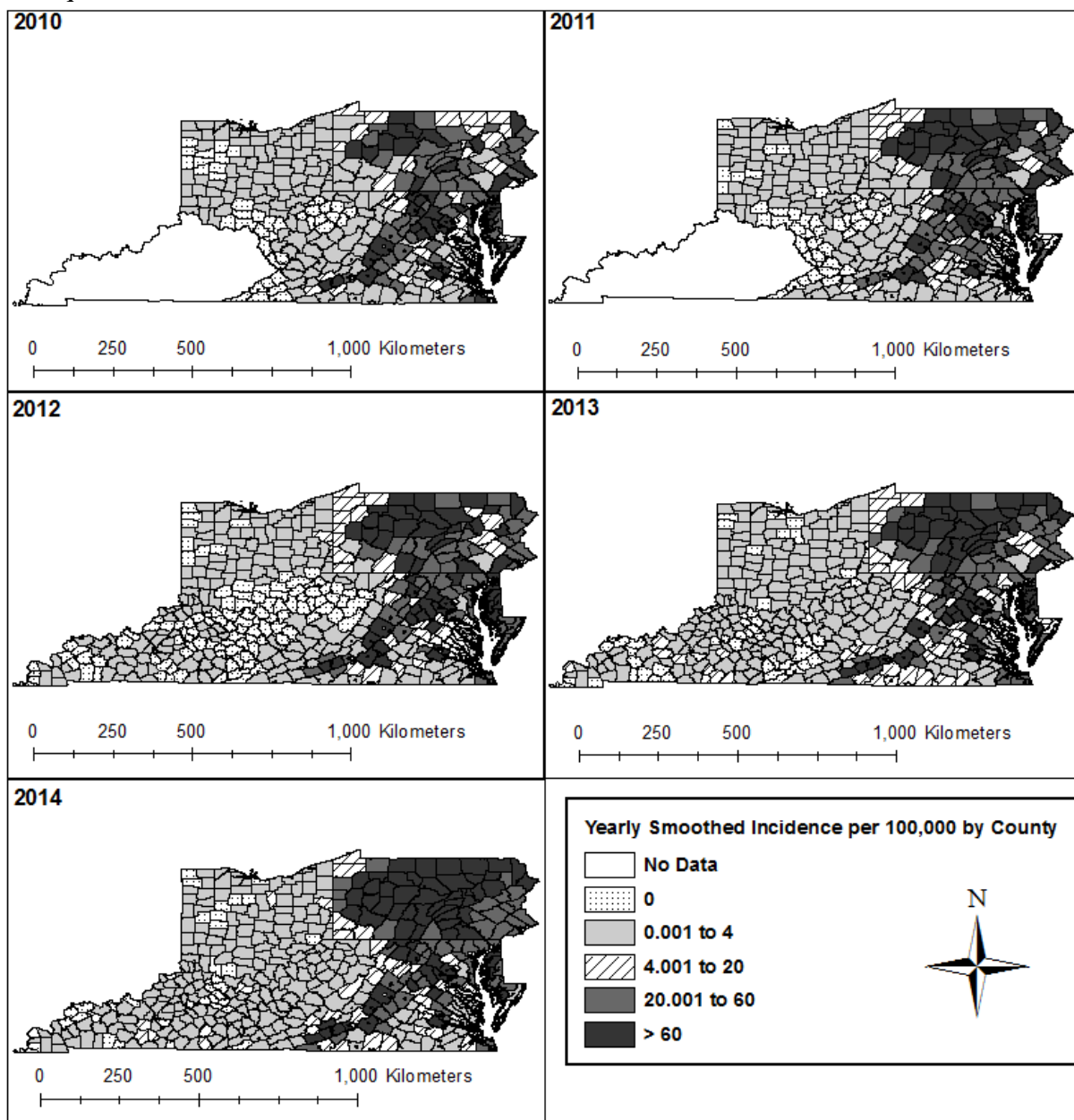


Figure 1.4 Local Indicators of Spatial Autocorrelation (LISA) Cluster Analysis incorporating rook spatial weight, 99,999 permutations, and a pseudo p-value of 0.05 (Anselin 2003). Counties in dark grey are “hot spots” or high incidence counties surrounded by high incidence counties. Light grey counties are “cold spots” or low incidence counties surrounded by other low incidence counties. Counties in medium grey are statistically significant outliers or low incidence counties bordered by high incidence counties.

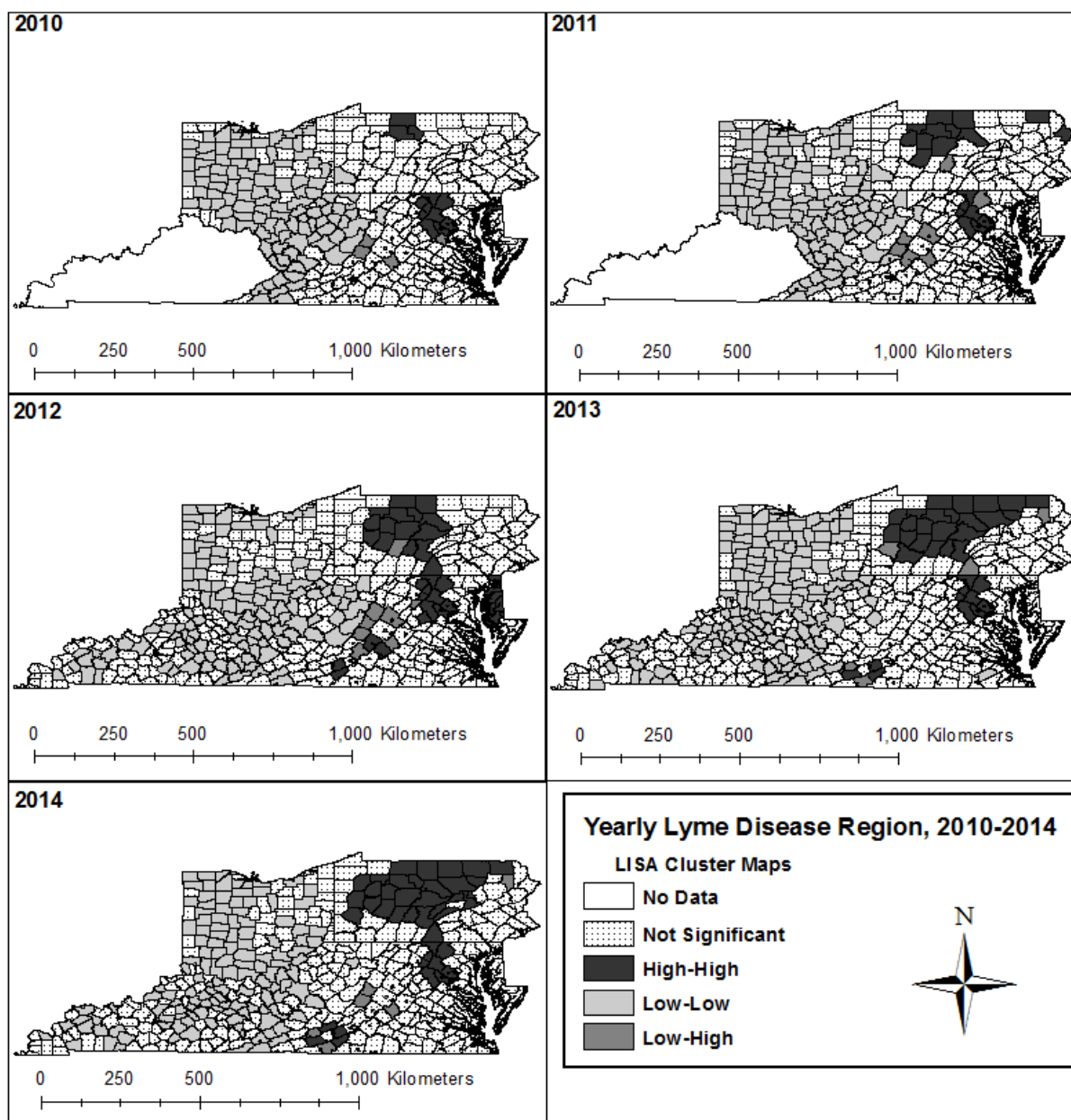
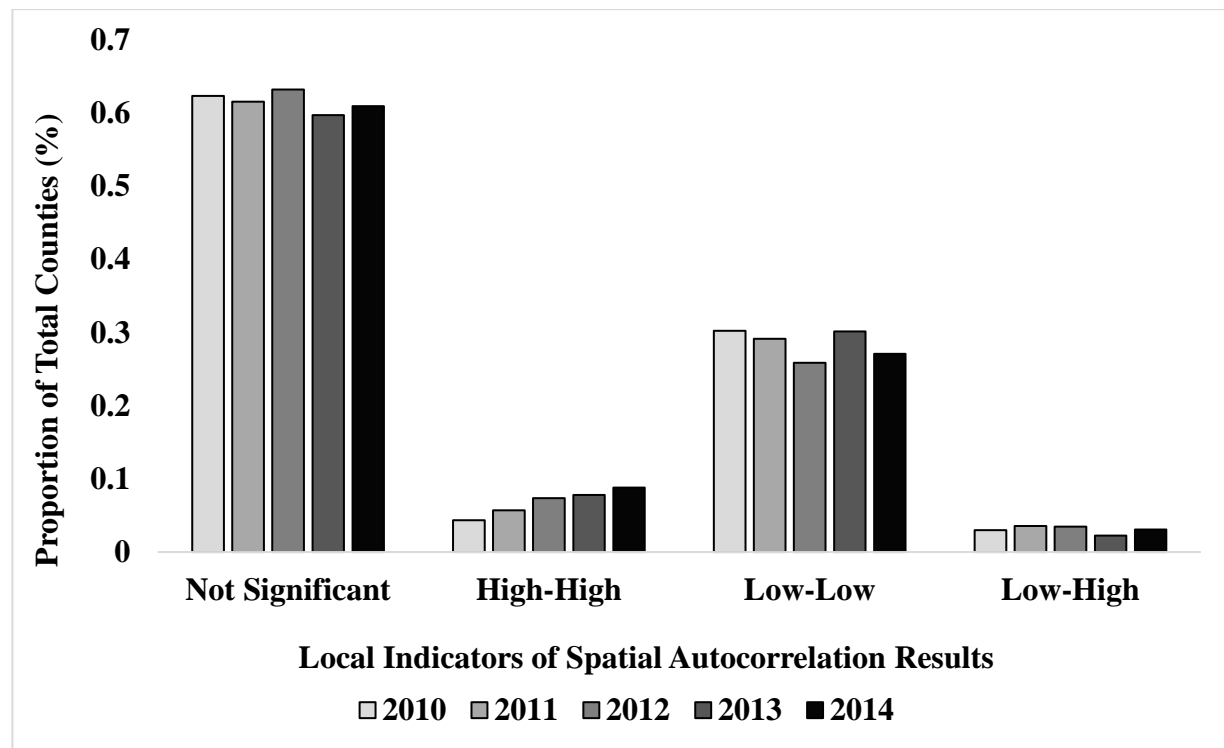


Figure 1.5 Graphical Representation of the Proportion of County Level Lyme Reporting Hotspots (High-High), Cold Spots (Low-Low), and Statistically Significant Spatial Outliers (Low-High) Within the Multistate Region, 2010-2014.



CHAPTER TWO

Using Exploratory Spatial Data Analysis to Identify and Characterize Patterns of Travel Associated Confirmed Lyme disease Cases in West Virginia, 2011-2015

ABSTRACT

Travel associated cases of Lyme disease are theorized to have a substantial effect on case visualization. Yet no study to date has formally focused on identifying the impact of travel related cases on geographic visualization of Lyme disease occurrence and spread. This study's aims were to confirm travel as a source of geographic bias and 2) characterize spatial patterns associated with the proportion of cases attributable to travel using West Virginia Public health Lyme disease surveillance data. Confirmed cases were obtained for 2011-2015 from the West Virginia Bureau of Public Health. Following data preprocessing, variables representing total overall confirmed cases and only those cases obtaining infection within county of residence were visualized in separate maps and analyzed using a paired t-test. Spatial autocorrelation was assessed for the proportion of cases attributable to travel using a Local Moran's I with empirical Bayesian rate. The paired t-test indicated statistically significant ($t = 3.99$, $df = 54$, $P = 0.0002$) non-equivalency between total overall cases and those obtained within patient's home county. Local Moran's I identified statistically significant autocorrelation ($I = 0.17$, pseudo $P = 0.02$) and both clusters and spatial outliers. Findings suggest that travel associated case acquisition is a significant source of geographic bias in visualization of Lyme disease cases.

Keywords: Travel Associated Case, Lyme disease, Surveillance

INTRODUCTION

Lyme disease caused by human exposure to the bacterial agent *Borrelia burgdorferi*, is the most common vectorborne disease in the United States (1, 2). Human infection results in multi-symptom conditions and can lead to serious health consequences if left untreated (3). While cases have been documented within each state of the contiguous U.S. (4), a proportion are regarded as travel-associated. The concept of travel associated Lyme disease cases was noted as early as 1979 (5) and has since been hypothesized as a source of confounding for case visualization (6, 7). The potential for confounding is exacerbated in low incidence areas where up to 80% of cases are potentially attributable to recent travel to high incidence areas (8). Yet, no study to date has confirmed that travel-associated Lyme disease cases is a confounding factor for spatial analysis within public health surveillance data. Consequently, this retrospective investigation aims to 1) establish travel as statistically significant source of geographical bias in visualization of Lyme disease surveillance data and 2) identify spatial patterns in the proportion of travel related cases reported in West Virginia.

METHODS

Confirmed human Lyme disease case counts for 2011-2015 were provided by the West Virginia Bureau of Public Health (9). This period was chosen because of similarities in Lyme disease case definitions, providing consistency among epidemiologic, medical, and laboratory evidence (10). The data contains variables for the patient's home county, the reporting year, and a dummy variable indicating whether a case was travel-related. Criteria used by local health departments to designate cases as travel related are based upon the patient's self-reported travel history and most likely location of tick exposure (11). Three county-level (n=55) variables were created using the travel-related disease attribute, representing total cases, cases infected within the patient's home

county, and proportion of cases attributable to travel. Spreadsheets were joined in ArcMap (ESRI, Redlands CA) to a United States Counties and Localities shape file (12). Choropleth maps were created for total cases and those obtaining infection within their county of residence. County-level choropleth class breaks were set as 0, 1 to 4, 5 to 20, 21 to 60, and > 60 cases to place specific emphasis on counties with low to moderate occurrence of Lyme disease (6). A paired t-test and Wilcoxon Signed Rank test were conducted utilizing SAS 9.4 (SAS Institute, Cary NC) to determine if the county level differences between total overall confirmed Lyme disease cases and confirmed cases obtained within patient's home county were statistically equivalent to zero. Univariate Local Moran's I analyses with empirical Bayesian rate were conducted in GeoDa 1.8 (GeoDa Center, Tempe AZ) to identify spatial patterns, and adjust for instability of variances in rates calculated (13, 14). A queen's contiguity matrix was used to identify neighbors in the local analysis, and results were permuted 99,999 times to detect statistically significant clustering and outliers at a pseudo p -value of 0.05 (15).

RESULTS

There were 660 confirmed Lyme disease cases in West Virginia from 2011-2015, of which 541 (82%) were obtained within the patient's home county. Statistically significant non-equivalency was identified in paired t-test ($t = 3.99$, $df = 54$, $P = 0.0002$) and Wilcoxon Signed Rank test ($S = 264$, $P < 0.001$) among variables representing total overall confirmed cases of Lyme disease and those which obtained infection within their home county. The number of counties within each of the choropleth classes differed between inclusion or exclusion of travel related cases (Figure 2.1). Univariate Local Moran's I with empirical Bayesian rate detected statistically significant ($I = 0.17$, pseudo $P = 0.02$) clustering and spatial outliers (Figure 2.2). Counties ranked as high-high have a high proportion of cases attributable to travel and border counties which on average

also have a high proportion of travel related cases. Conversely, low-low and low-high counties have a low proportion of cases attributable to travel which border counties with a low (low-low) or high (low-high) average proportion of travel related cases.

DISCUSSION

Travel is commonly mentioned as a potential source of bias in Lyme disease literature. West Virginia is an ideal state to study geographic biases associated with Lyme disease case reporting as it is a transitional state, bordered on the east by high incidence states of Maryland, Virginia, and Pennsylvania and on the west by low incidence states of Kentucky and Ohio (16). This is the first known study to employ statistical and spatial analytic methodologies to confirm the necessity of adjusting for travel associated cases within public health surveillance data prior to geographic visualization. Adjustment for travel resulted in statewide changes in the number of counties represented within each of the choropleth classes (Figure 2.1). In particular, the choropleth classes with zero, 1 to 4, and 5 to 20 confirmed Lyme disease cases displayed dramatic ($\pm 16\text{-}50\%$) changes in the percent of counties represented after removing cases that were attributable to travel. Conversely, exclusion of travel related cases did not result in much variability ($\pm 1.8\%$) among the percentage of counties represented within the choropleth classes of 21 to 60 and > 60 confirmed Lyme disease cases. This suggests that counties that reported a low number of cases are particularly susceptible to travel related biases. Ultimately, these changes in the percent of counties within the case based choropleth classes resulted in visualization of two separate maps which were statistically non-equivalent ($t = 3.99$, $df = 54$, $P = 0.0002$) (Figure 2.1).

Spatial patterns in the proportion of cases attributable to travel were identified utilizing local indicators of spatial autocorrelation. The cluster map (Figure 2.2) displays significant county-

level clustering in the proportion of cases attributable to travel. The proportion of confirmed Lyme disease cases attributable to travel by county are provided as a label over LISA results to provide reference to the crude estimates. Counties in western West Virginia and the Eastern Panhandle display no significant clustering or clustering of low values of the proportion of cases attributable to travel. Alternatively, northern and south central West Virginia counties displayed clustering of high proportions with a few spatial outliers with low values neighboring higher values. It is important to note that while some counties experienced high proportion of travel related cases, not all were denoted as high-high in the LISA cluster map. This discrepancy resulted because the value in a given county was high, but was bordered by neighboring counties with low proportion of travel related cases (17).

The choropleth maps and patterns of LISA rankings are consistent with previous research regarding travel biases (8) and highlight potential spatial biases which influence consistency of non-travel related case reporting due to the combination of low and high incidence states which share a border with West Virginia (6). This spatial bias may be especially relevant in counties that form clusters of high proportions of travel-related cases, indicating at-risk areas where Lyme disease is not yet endemic, although there are enough travel-related cases that the disease may soon spread to these locations. Future research is warranted to incorporate study findings and examine other potential geographic biases related to public health surveillance data.

LITERATURE CITED

1. CDC. How Many People Get Lyme Disease. 2015. (<http://www.cdc.gov/lyme/stats/humancases.html>). (Accessed 8/11/2016).
2. Hinckley A, Connally N, Meek J, et al. Lyme Disease Testing by Large Commercial Laboratories in the United States. *Clinical Infectious Diseases* 2014;59(5):676-81.
3. Stanek G, Wormser G, Gray J, et al. Lyme borreliosis. *The Lancet* 2012;379:461-73.
4. CDC. Lyme disease data tables. 2015. (<http://www.cdc.gov/lyme/stats/tables.html>). (Accessed 8/11/2016).
5. Steere AC., Malawista SE. Cases of Lyme Disease in the United States: Location Correlated with Distribution of Ixodes dammini. *Annals of Internal Medicine* 1979;91:730-3.
6. Hendricks B., Mark-Carew M. using exploratory data analysis to identify and predict patterns of human Lyme disease case clustering within a multistate region, 2010-2014. *Spatial and Spatio-temporal Epidemiology* 2017;20:35-43.
7. Szonyi B, Srinath I, Esteve-Gassent M, et al. Exploratory spatial analysis of Lyme disease in Texas -what can we learn from the reported cases? *BMC Public Health* 2015;15(924):1-8.
8. Forrester J, Brett M, Matthias J, et al. Epidemiology of Lyme disease in low-incidence states. *Ticks and Tick-borne Diseases* 2015;6:721-23.
9. West Virginia Bureau of Public Health. 2015 Confirmed County-Level Human Lyme disease case surveillance data. 2016.
10. Li J, Hong K, Duan Y, et al. Spatial and Temporal Emergence Pattern of Lyme Disease in Virginia. *American Journal Tropical Medicine and Hygiene* 2014;91(6):1166-72.
11. West Virginia Bureau of Public Health. Lyme Disease Case Report Form. 2017. (dhhr.wv.gov/oeps/disease/WVEDSS/Documents/lyme.pdf). (Accessed 4/12/2017).
12. USDA NRCS. Geospatial Gateway. 2015. (<https://gdg.sc.egov.usda.gov/>). (Accessed 1/20/2016).
13. Assuncao RM., Reis EA. A New Proposal to Adjust Moran's I for Population Density. *Statistics in Medicine* 1999;18:2147-62.
14. Anselin L, Syabri I, Kho Y. GeoDa: An Introduction to Spatial Data Analysis. *Geospatial Analysis* 2006;38:5-22.
15. Anselin L. GeoDa 0.9 user's Guide. In: Illinois Uo, ed. *Center for Spatially Integrated Social Science*, 2003.
16. CDC. Lyme disease data and statistics. 2015. (<https://www.cdc.gov/lyme/stats/index.html>). (Accessed 2/11/2016).
17. Burt JE., Barber GM., DL. R. *Elementary Statistics for Geographers*. New York, NY: The Guilford Press; 2009.

Figure 2.1 Graphical Representation of the Overall Total and Total Obtained within Patients Home County Confirmed Human Lyme disease Cases in West Virginia, 2011-2015.

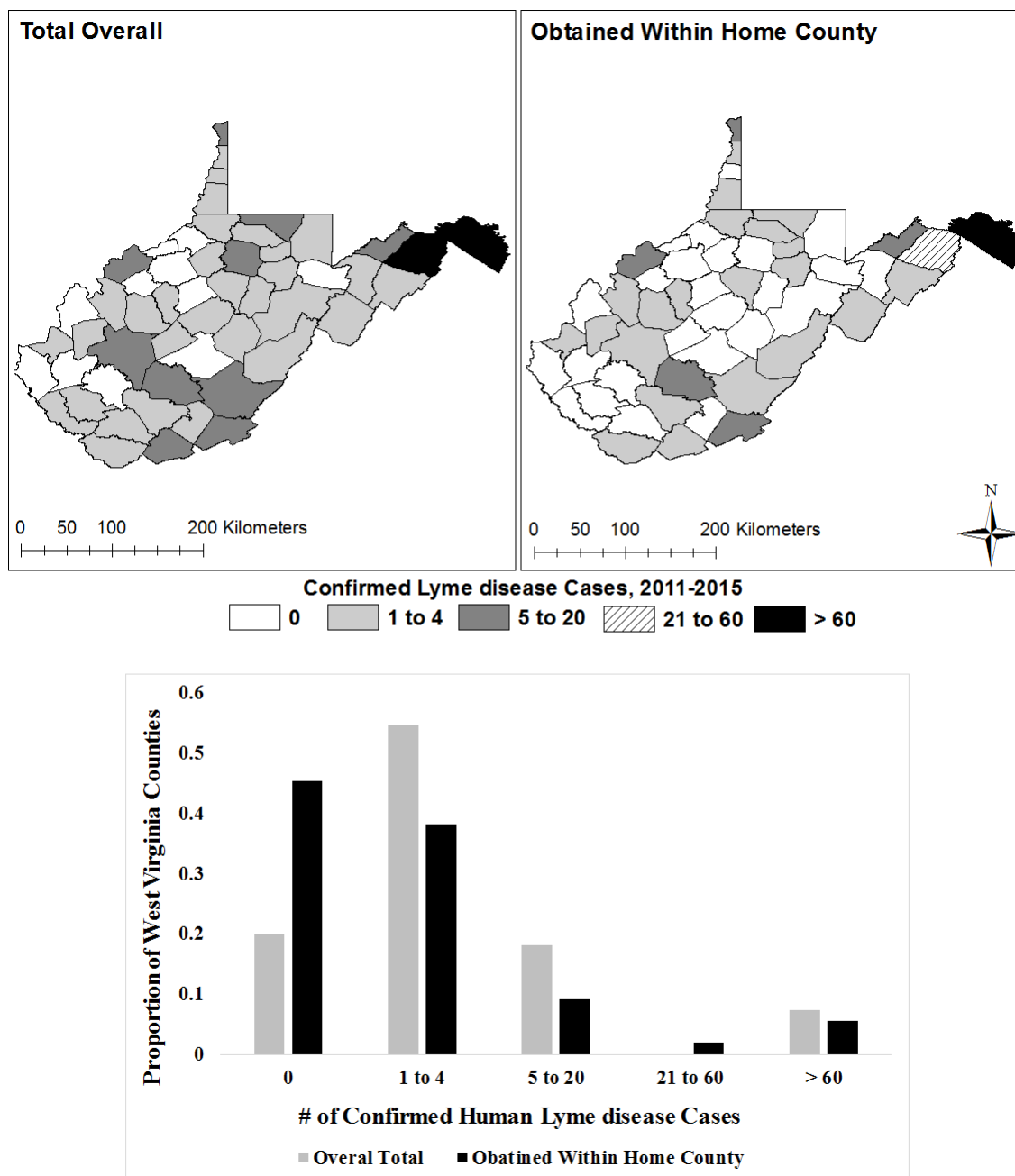
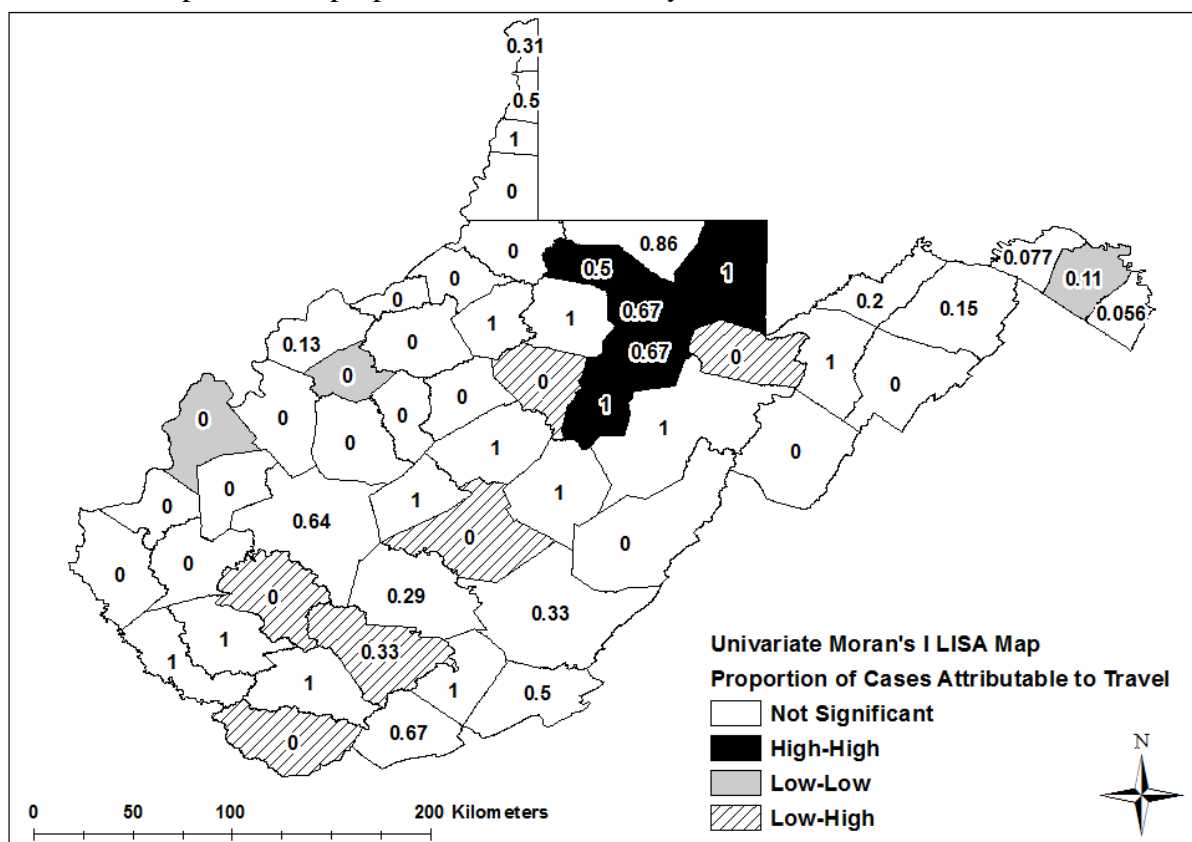


Figure 2.2 Univariate Local Moran's I with Empirical Bayesian Rate. Numbers within each of the counties represent the proportion of confirmed Lyme disease cases attributable to travel.



CHAPTER THREE

Using Geo-Analytic Techniques to Evaluate the Utility of Companion Animal Tick Surveillance Practices for Monitoring Spread and Occurrence of Human Lyme disease in West Virginia, 2014-2016

ABSTRACT

Companion animals such as dogs and cats are potentially effective sentinel populations for monitoring occurrence and spread of Lyme disease. No study to date has evaluated the public health utility of sentinel programs using geo-analytic approaches despite the deep-rooted spatial processes that influence tick-host disease transmission. Confirmed Lyme disease cases in West Virginia and *Ixodes scapularis* submitted by veterinarians to the state health department were obtained for 2014-2016. Separate ordinary least squares and spatial lag regression models were conducted to estimate the association between average number of *Ixodes scapularis* collected on pets and human Lyme disease incidence per 1,000 persons. Residuals from regression models were visualized using Local Moran's I as a diagnostic tool to identify spatial dependence among regression models. Statistically significant associations were identified between average numbers of *Ix. scapularis* collected from dogs and human Lyme disease in the OLS ($\beta=0.20$ $P < 0.001$) and spatial lag ($\beta = 0.12$, $P = 0.002$) regression. No significant associations were identified for cats in either regression model. Statistically significant ($P \leq 0.05$) spatial dependence was identified in all regression models. Local Moran's I maps produced for spatial lag regression residuals indicated a decrease in model over and underestimation, but identified a higher number of statistically significant outliers than OLS regression. Results suggest that dogs are effective sentinel populations for monitoring human Lyme disease infection. Findings identify significant spatial effects associated with regression residuals, emphasizing the

importance of statistical methodologies, which can accommodate the spatial structure of public health surveillance data.

Keywords: Lyme disease, Surveillance, Spatial Analysis, Companion Animals

INTRODUCTION

Lyme disease is the most commonly reported vectorborne disease in the United States (1, 2). Human infection occurs primarily in the Northeastern and North Central United States, but is not exclusive to these areas due to constant geographic expansion of the principal vector *Ixodes scapularis* (3, 4). Human risk of exposure is uncertain in many areas where Lyme disease is considered an emerging issue due to variability in recognition of infection and disease manifestation reported by health officials, biases among diagnostic tests, and insensitivity in local tick surveys at early stage invasion (5-7). Tick surveys among companion animals have been identified as a potentially representative surveillance methodology for estimating human risk within geographic areas where Lyme disease is an emerging concern (8). Animals such as dogs, cats, and horses are potentially effective sentinel populations due to increased likelihood of tick infestation and close association with their human owners (9). Few studies to date have evaluated the utility of companion animal sentinel programs as public health tools with spatial analytic approaches to study the spatial structure of tick-host disease transmission dynamics (10).

Humans and companion animals are dead end hosts for *Borrelia burgdorferi*, causative agent of Lyme disease (11, 12). Exposure to tickborne pathogens occurs when a host comes in contact with actively questing ticks in the environment. Presence of infected ticks are influenced by environmental factors such as availability of competent reservoir hosts and climatic factors that drive tick survival and pathogen propagation (13, 14). Domestic pets, such as dogs, are significant risk factors for disease acquisition, and six times more likely to be exposed to infected ticks due to the increased potential exposure time in tick habitat (7, 15). Sampling companion animal populations is an effective approach to estimate human risk in areas where Lyme disease may be newly endemic.

Popular concern for Lyme disease has prompted many animal sentinel studies across the United States and the United Kingdom (6, 7, 16-19). These studies primarily focused on ticks and associated pathogens collected from domestic dogs because of their ability to produce antibodies to *Borrelia burgdorferi*, attainable travel history information, and frequency of outdoor exposure (6, 11, 20-24). Methodologies for pathogen detection are restricted to serologic tests such as enzyme-linked immunosorbent assay and Western blot, as well as direct polymerase chain reaction (PCR) of ticks collected from domestic dogs (6, 7, 16-19). Studies suggest that serology screening is an effective approach in Lyme endemic areas (25, 26) but may not be as sensitive an indicator for estimating prevalence of infected ticks as direct screening of ticks via PCR within areas of low tick density (7). In addition to potential screening biases, previous studies also suggest variability in sample size by locality may have influenced study findings (19).

Some studies have shown that success in detection of *Borrelia burgdorferi* in dogs or ticks collected from dogs correlates with human Lyme disease incidence and tick feeding behavior within their respective localities (11, 27). This conclusion, along with recognition that tick survival and disease transmission are spatial processes (13), recommend analytic spatial approaches to identify areas of increased human risk. To date, no known study has applied analytic spatial techniques to investigate geographic associations between *Ixodes scapularis* collected from sentinel companion animal surveillance and human Lyme disease cases, investigated sentinel potential for companion animals other than dogs, or employed spatial tools to identify areas where increased effort may be needed to optimize Lyme disease surveillance efforts. The objectives of our study are to 1) conduct spatial regression to identify geographic associations between animal species-specific *Ixodes scapularis* sentinel and human Lyme disease

case data and to 2) investigate local clustering of regression residuals to identify counties where additional effort may be needed to optimize surveillance efficacy.

METHODS

Study Area

Annual confirmed human case counts within West Virginia have increased two-fold between 2005 (n=61) and 2014 (n=136) (3). Geographically, West Virginia is bordered on the east by high incidence states of Pennsylvania, Maryland, and Virginia and on the west by the low incidence states of Ohio and Kentucky (CDC 2015). In addition to human cases reported, companion animal sentinel surveillance data has been collected from local licensed veterinary practices within the state since 2013 (28).

Data acquisition and management

Human Lyme disease case data in West Virginia were obtained for 2014-2016 from the West Virginia state health department (29). Human surveillance data were limited to confirmed cases (based on the 2011 Council of State and Territorial Epidemiologists Lyme disease case definition) with reported infections acquired within their home county. Restricting raw case counts limited biases associated with travel related disease acquisition (30), led to consistency of medical, laboratory, and epidemiologic evidence (31), and included recent exposure to tick habitat (32). Crude measures of human Lyme disease frequency were calculated by dividing the sum of 2014-2016 county-level case counts by 2015 U.S. census West Virginia county-level population estimates (33) to reflect county-level incidence throughout the state from 2014-2016 (34).

Companion animal surveillance data obtained from the West Virginia state health department from 2014-2016 and contained yearly county level information relating the diversity and abundance of different tick species removed from animals by veterinarians participating in the convenience sample. Sentinel surveillance data requested were limited to 2014-2016, because of improved data quality following the initial year of data collection. Observations were sorted by county, and *Ixodes scapularis* removed from animals were summed across individuals in SAS 9.4.1 (SAS Institute, Cary NC) to produce four variables representing the total and average number of *Ixodes scapularis* collected from dogs and cats by county. All licensed veterinary practices within West Virginia received equal opportunity for voluntary participation. Therefore, zeros were assigned to any county that had no *Ixodes scapularis* collection data within our study period.

Human population and case estimates were paired with average number of *Ixodes scapularis* estimates from dogs and cats by respective reporting county in Microsoft Excel. Spreadsheet were joined in ArcMap 10.4.1 (ESRI Redlands, CA) to a U.S. Counties and Localities shape file obtained U.S. Department of Agriculture National Resources Conservation Service Geospatial Gateway (USDA:NRCS) (35).

Regression and Spatial Analysis

Separate ordinary least squares regression models were performed using average number of *Ixodes scapularis* collected on dogs and cats as the independent variable and county level human Lyme disease incidence per 1,000 as the dependent variable. Rook and queen spatial contiguity weights matrices were visualized within connectivity histograms to characterize neighbor structures (36). Contiguity based spatial weights were utilized within our spatial analysis to satisfy regularity conditions and reduce the potential for introducing heteroscedasticity (37). A

queen contiguity weight was chosen to increase the number of neighbors evaluated during the spatial analysis (38). Univariate Local Moran's I and LISA maps were calculated to identify the extent of spatial autocorrelation within the residuals (39). Lagrange multiplier values were calculated during the OLS regression utilizing the previously defined queen spatial weights matrix, as a model selection tool in comparing spatial lag and error regression models (38, 40). Residuals from the spatial model were visualized using univariate Local Moran's I LISA maps to display remaining spatial dependence. Akaike information criterion between OLS and spatial autoregressive models were compared to examine model goodness of fit. All regression and spatial analyses were conducted in GeoDa 1.8 (GeoDa Center Tempe, AZ).

RESULTS

Regression and Spatial Analysis

County-level incidence of confirmed human Lyme disease ranged between 0-3.88 per 1,000 persons with an average rate of 0.30 (± 0.65) cases per 1,000 persons. County-level *Ixodes scapularis* collections ranged from 0-93 for dogs with a mean number removed of 11.2 (± 18.8) and 0-37 for cats with a mean number of 5.6 (± 9.30). Ordinary least squares regression parameters were significant ($F = 3.91$, $P < 0.001$) for the dog specific model, and not significant ($F = 0.91$, $P = 0.36$) for the cat specific model. Univariate Local Moran's I indicated significant clustering among OLS canine ($I = 0.27$, $pseudo P = 0.002$) and cat ($I = 0.50$, $pseudo P < 0.001$) regression residuals. Lagrange multiplier (LM) values were significant for dog spatial lag ($F = 24.0$, $P < 0.01$) and dog spatial error ($F = 8.19$, $P = 0.04$) as well as cat spatial lag ($F = 30.1$, $P < 0.01$) and cat spatial error ($F = 27.2$, $P < 0.001$) models. Spatial lag autoregressive model was chosen based upon higher robust LM values (30.0 vs. 14.0 for dogs and cats (12.8 vs. 9.92) (38), and because the spatial lag model adjusts for substantive rather than nuisance spatial dependence

(40). Average number of *Ixodes scapularis* collected per dog remained significant ($F = 3.04$, $P = 0.002$) in the spatial lag model, and identified significant positive spatial dependence ($\rho = 0.66$, $P < 0.001$). Average number of *Ixodes scapularis* collected per cat remained non-significant ($F = -0.01$, $P = 0.88$) within the spatial lag model. However, a significant ($\rho = 0.73$, $P < 0.001$) strong positive spatial dependence was identified. Univariate Local Moran's I indicated significant dispersal among canine ($I = -0.14$ *pseudo* $P = 0.02$) and non-significant dispersal among feline ($I = -0.04$ *pseudo* $P = 0.40$) spatial lag residuals. Parameter estimates for animal specific regression models are summarized in Table 3.1. Model goodness of fit, measured by a reduction in Akaike Information Criterion, was improved for dog (AIC = 76.2 vs. AIC = 98.5) and cat (AIC = 84.5 vs. 111.5) models utilizing spatial lag regression. Local indicators of spatial autocorrelation (LISA) maps were produced for OLS and spatial lag model residuals for dogs and cats (Figure 3.1). High-high ranking indicates counties where human Lyme disease cases per 1,000 persons are higher than what would be expected after adjusting for covariates. Conversely, low, high-low, and low-high rankings indicate counties where human Lyme disease cases per 1,000 are lower, higher surrounded by lower, or lower surrounded by higher than what would be expected after adjusting for covariates.

DISCUSSION

Exploratory statistical approaches are routinely applied as methods to evaluate associations between animal sentinel and human disease data. This is the first known study to assess associations between sentinel surveillance data from dogs and cats and human Lyme disease incidence utilizing a combination of ordinary least squares and spatial regression models. Significant associations between *Ix. scapularis* collections from dogs were identified in OLS and spatial lag models. This finding is consistent with previous studies which identified dogs as

potentially effective sentinel animals for monitoring human Lyme disease due to their close association with pet owners, and increased probability for tick exposure (7, 9, 15). Conversely, no statistical association was identified between feline *Ix. scapularis* collections and human disease in either OLS or spatial lag autoregressive models. While this may seem unexpected due to the similarities in shared exposure risk between domestic cats and dogs, prevalence and manifestation of Lyme disease in felines is poorly understood (41).

Further analysis of model residuals via local indicators of spatial autocorrelation permitted additional insight into surveillance system effectiveness by providing a sound diagnostic methodology for identifying spatial dependence within the regression models (39, 42). Spatial dependence, also termed spatial autocorrelation, results when the assumption of independence among dependent variables is violated, or when the regression residuals themselves display spatial dependence (42). In either case, parameter estimates from OLS regression are inefficient. Local indicators of spatial autocorrelation (LISA) (Figure 1) calculated for OLS regression residuals identified significant clustering of model over and underestimation throughout West Virginia. Incorporation of the spatial lag model resulted in a reduction in the number of counties regarded as high-high (model underestimation) and low-low (model overestimation) among dogs and cats. However, residual spatial dependence was identified within spatial lag model residuals in the form of low-high and high-low spatial outliers. These spatial outliers are potentially an artifact of border bias (36), resulting from a shared border with the Lyme disease endemic states of Pennsylvania, Maryland, and Virginia (1). Additionally, they could also be spurious outliers resulting from low veterinary participation, and potentially remedied by increasing veterinary practice recruitment efforts in West Virginia northeastern panhandle counties.

Study findings are based upon a statewide county level ecological regression between human Lyme disease incidence per 1,000 persons and animal sentinel data collected as a convenience sample from 2014-2016. Spatial lag models are often conducted for data aggregated at the county level, but are potentially prone to ecological fallacy if data incorporated are at differing spatial scales (37). Additionally, some studies call for the presentation of vector-borne disease surveillance data at scales finer than the county level (43). We attempt to adjust for ecological fallacy and misclassification biases by incorporating confirmed cases of human Lyme disease obtained within patient's home county of residence and animal sentinel data aggregated at the same county level scale. Our analysis was limited to county-level scale as this is the finest scale for which travel-associated infection status is noted. Other potential limitations are associated with differing reporting biases associated with unequal Lyme disease awareness among physicians or other public health officials, and varying presentation of Lyme disease signs and symptoms among patients infected (44-46). Unfortunately, we are unable to characterize the effect of differing reporting biases due to the aggregate nature of our retrospective investigation.

Despite these limitations, findings fill significant gaps in the literature regarding the spatial biases of public health surveillance data. Results confirm the efficacy of companion animal sentinel surveillance programs by further implicating dogs as risk factors for human Lyme disease infection. Interestingly, residual spatial dependence was identified among spatial models and potentially point to other geographic biases not considered. In particular, West Virginia is prone to border biases resulting from shared boundaries with both high and low incidence states (1, 36). Results highlight the unique geographic position that West Virginia shares with its neighboring states, and emphasize the necessity of spatial analysis to adjust for inherent biases associated with collection of public health surveillance data.

LITERATURE CITED

1. CDC. How Many People Get Lyme Disease. 2015.
(<http://www.cdc.gov/lyme/stats/humancases.html>). (Accessed 8/11/2016).
2. Hinckley A, Connally N, Meek J, et al. Lyme Disease Testing by Large Commercial Laboratories in the United States. *Clinical Infectious Diseases* 2014;59(5):676-81.
3. CDC. Lyme disease data and statistics. 2015.
(<https://www.cdc.gov/lyme/stats/index.html>). (Accessed 2/11/2016).
4. Mead P.S. Epidemiology of Lyme disease. *infect Dis Clin N Am* 2015;29:187-210.
5. GUERRA MA., WALKER ED., U. K. CANINE SURVEILLANCE SYSTEM FOR LYME BORRELIOSIS IN WISCONSIN AND NORTHERN ILLINOIS: GEOGRAPHIC DISTRIBUTION AND RISK FACTOR ANALYSIS [electronic article]. *Am J Trop Med Hyg*.
6. Duncan Aw., Correa MT., Levine JF., et al. The Dog as a Sentinel for Human Infection: Prevalence of *Borrelia burgdorferi* C6 Antibodies in Dogs from Southeastern and Mid-Atlantic States. *Vector-Borne and Zoonotic Diseases* 2005;5(2):101-9.
7. Hamer SA., Tsao JL., Walker ED., et al. Use of tick surveys and serosurveys to evaluate pet dogs as a sentinel species for emerging Lyme disease. *AJVR* 2009;70(1):49-56.
8. Millen K., Kugeler KJ., Hinckley AF., et al. Elevated Lyme Disease Seroprevalence Among Dogs in a Nonendemic County: Harbinger or Artifact. *Vector-Borne and Zoonotic Disease* 2013;13(5):340-1.
9. Anderson JF. Epizootiology of *Borrelia* in Tick Vectors and Reservoir Hosts. *Reviews of Infectious Diseases* 1989;11(6):1451-9.
10. Killilea M, Swei A, Lane R, et al. Spatial Dynamics of Lyme Disease: A Review. *EcoHealth* 2008;5:167-95.
11. Little SE., Heise SR., Blagburn BL., et al. Lyme Borreliosis in dogs and humans in the USA. *Trends in Parasitology* 2010;26(4):213-8.
12. Radolf JD, Caimano, M.J., Stevenson, B., Hu, L.T.,. Of ticks, mice and men: understanding the dual-host lifestyle of Lyme disease spirochaetes [electronic article]. *Nat Rev Microbiol*.
13. Eisen RJ., Piesman J., Zielinski-Gutierrez E., et al. What Do We Need to Know About Disease Ecology to Prevent Lyme Disease in the Northeastern United States? *Journal of Medical Entomology* 2012;49(1):11-22.
14. Brinkerhoff J, Gilliam W, Daines D. Lyme Disease, Virginia, USA, 2000-2011. *Emerging Infectious Diseases* 2015;20(10):1661-68.
15. Jones TF., Garman RL., LaFluer B., et al. Risk factors for Tick Exposure and Suboptimal Adherence to Preventive Recommendations. *American Journal of Preventive Medicine* 2002;23(1):47-50.
16. Smith FD., Ballantyne R., Morgan ER., et al. Estimating Lyme disease risk using pet dogs as sentinels. *Comparative Immunology, Microbiology, and Infectious Diseases* 2012;35:163-7.
17. Goosens HAT., Van Den Bogard AE., MKE. N. Dogs as Sentinels for Human Lyme Borreliosis in The Netherlands. *Journal of Clinical Microbiology* 2001;39(3):844-8.
18. Glickman LT., Moore GE., Glickman NW., et al. Purdue University-Banfield National Companion Animal Surveillance Program for Emerging and Zoonotic Diseases. *Vector-Borne and Zoonotic Diseases* 2006;6(1):14-23.

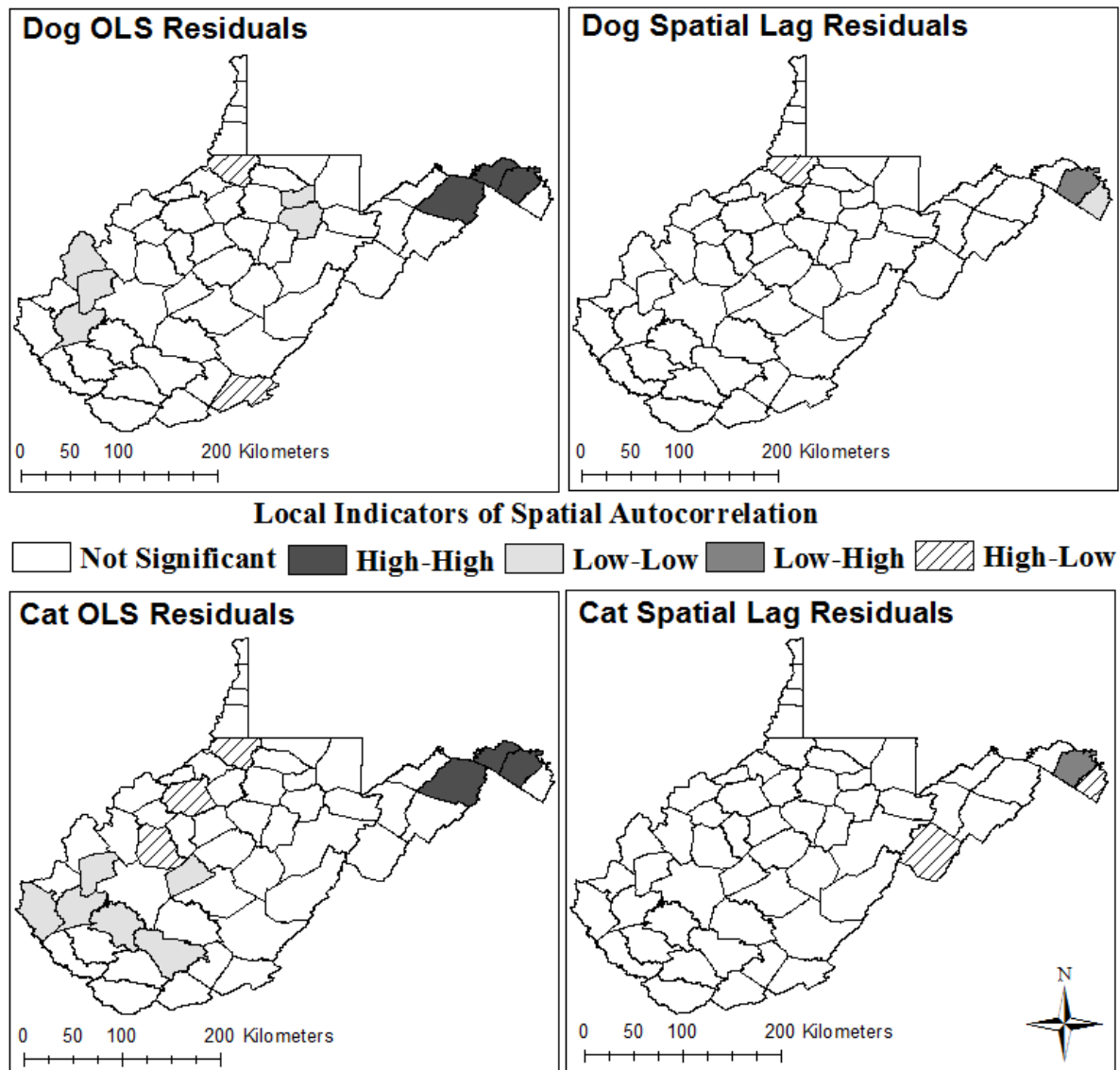
19. Johnson JL., Ginsberg HS., Zhioua E., et al. Passive Tick Surveillance, Dog Seropositivity, and Incidence of Human Lyme disease. *Vector-Borne and Zoonotic Diseases* 2004;4(2):137-42.
20. Mather TN., Fish D., RT. C. Competence of dogs as resevoirs for Lyme disease spirochetes (*Borrelia burgdorferi*). *J AM Vet Med Assoc* 1994;15:186-8.
21. Burgess EC. Natural exposure of Wisconsin dogs to the Lyme disease spirochete (*Borrelia burgdorferi*). *Lab Anim Sci* 1986;36:288-90.
22. Walker ED., Stobierski MG., Poplar ML., et al. Geographic distribution of ticks (Acari:Ixodidae) in Michigan, with emphasis on *Ixodes scapularis* and *Borrelia burgdorferi*. *Journal of Medical Entomology* 1998;35(5):872-82.
23. Eng TR., Wilson ML., Spielman A., et al. Greater risk of *Borrelia burgdorferi* infection in dogs than in people. *J Infect Dis* 1988;158:1410-1.
24. Lindenmayer JM., Marshall D., Onderdonk AB. Dogs as sentinels for Lyme disease in Massechussetts. *Am J Public Health* 1991;81:1448-55.
25. Bowman D., Little SE., Lorentzen L., et al. Prevalence and geographic distribution of *Dirofilaria immitis*, *Borrelia burgdorferi*, *Ehrlichia canis*, and *Anaplasma phagocytophilum* in dogs in the United States: results of a national clinic-based serologic survey. *Vet Parasitol* 2009;160:138-48.
26. Mead Paul GR, Kugler Kiersten. Canine Serology as Adjunct to Human Lyme Disease Surveillance [electronic article]. *Emerging Infectious Diseases*.
27. Eisen R, Lane R, Fritz C, et al. Spatial Patterns of Lyme Disease Risk in California Based On Disease Incidence Data and Modeling of Vector-Tick Exposure. *American Journal of Tropical Medicine and Hygiene* 2006;75(4):669-76.
28. West Virginia Bureau of Public Health. 2014-2016 West Virginia Veterinary Tick Submission Program Data. 2017.
29. West Virginia Bureau of Public Health. 2015 Confirmed County-Level Human Lyme disease case surveillance data. 2016.
30. Szonyi B, Srinath I, Esteve-Gassent M, et al. Exploratory spatial analysis of Lyme disease in Texas -what can we learn from the reported cases? *BMC Public Health* 2015;15(924):1-8.
31. Li J, Hong K, Duan Y, et al. Spatial and Temporal Emergence Pattern of Lyme Disease in Virginia. *American Journal Tropical Medicine and Hygiene* 2014;91(6):1166-72.
32. Centers for Disease Control and Prevention (CDC). Lyme Disease Case Ascertainment Criteria. 2017. (<https://wwwn.cdc.gov/nndss/conditions/lyme-disease/case-definition/2017/>). (Accessed April 2, 2017).
33. United States Census Bureau. Population Estimates. 2015. (<http://www.census.gov/popest/>). (Accessed 1/20/2016).
34. Gertsman BB. *Epidemiology Kept Simple: An Introduction to Traditional and Modern Epidemiology*. West Sussex, UK: John Wiley & Sons; 2013.
35. USDA NRCS. Geospatial Gateway. 2015. (<https://gdg.sc.egov.usda.gov/>). (Accessed 1/20/2016).
36. Hendricks B., Mark-Carew M. using exploratory data analysis to identify and predict patterns of human Lyme disease case clustering within a multistate region, 2010-2014. *Spatial and Spatio-temporal Epidemiology* 2017;20:35-43.
37. Anselin L. Under the hood: Issues in the specification and interpretation of spatial regression models. *Agricultural Economics* 2002;27:247-67.

38. Anselin L, Syabri I, Kho Y. GeoDa: An Introduction to Spatial Data Analysis. *Geospatial Analysis* 2006;38:5-22.
39. Tiefelsdorf M. *Modelling Spatial Processes: The Identification and Analysis of Spatial Relationships in Regression Residuals by Means of Moran's I*.: Springer-Verlag; 2000.
40. Matthews SA. GeoDa and Spatial Regression Modelling. Presented at Penn State Population Research Institute 2006.
41. Krupka I., Straubinger RK. Lyme Borreliosis in dogs and cats: background, diagnosis, treatment, and prevention of infections with *Borrelia burgdorferi sensu stricto*. *Vet Clin North Am Small Anim Pract* 2010;40(6):1103-19.
42. Burt JE., Barber GM., DL. R. *Elementary Statistics for Geographers*. New York, NY: The Guilford Press; 2009.
43. Eisen L. ER. Need for Improved Methods to Collect and Present Spatial Epidemiologic Data for Vectorborne Diseases. *Emerging Infectious Diseases* 2007;13(12):1816-20.
44. Hayes E, Piesman J. How Can We Prevent Lyme Disease. *The New England Journal of Medicine* 2003;348:2424-30.
45. Bacon R, Kugeler K, Mead P. Surveillance for Lyme Disease --- United States, 1992--2006. *MMWR* 2008;57(SS10):1-9.
46. Brett M, Hinckley A, Zielinski-Gutierrez E, et al. U.S. healthcare providers' experience with Lyme and other tick-borne diseases. *Ticks and Tick-borne Diseases* 2014;5:404-8.

Table 3.1 Summary of Animal Species Specific Regression Results

Variable	Estimate	Std. Error	P-Value
Dog Specific Regression Models			
Model 1: Ordinary Least Squares Regression			
Average <i>Ix. scapularis</i> collected	0.20	0.06	< 0.001
Model 2: Spatial Lag Regression			
Average <i>Ix. scapularis</i> collected	0.12	0.04	0.002
Average Incidence per 1,000 Persons of Neighboring Counties	0.66	0.10	< 0.001
Cat Specific Regression Models			
Model 1: Ordinary Least Squares Regression			
Average <i>Ix. scapularis</i> collected	0.07	0.08	0.36
Model 2: Spatial Lag Regression			
Average <i>Ix. scapularis</i> collected	-0.01	0.55	0.88
Average Incidence per 1,000 Persons of Neighboring Counties	0.73	0.09	< 0.001

Figure 3.1 Local indicators of spatial autocorrelation maps output from univariate Local Moran's I for dog and cat Ordinary Least Squares and Spatial Lag Regression Residuals



PROJECT CONCLUSIONS

Mapping of surveillance data has become an increasingly popular method for relaying public health information due to the wide availability of geographic information systems. Approaches such as disease mapping and cluster analysis have proven useful in public health practice, particularly in reference to vectorborne disease (1, 2). Despite the recognition of place in human transmission cycles of vectorborne diseases, many studies have yet to characterize geographic biases affecting accurate representation of human cases. Our series of three studies were the first to identify the effects of boundary and travel related biases on human incidence of Lyme disease using ESDA approaches. We then adjust for these biases to increase accuracy of our regression models used to investigate associations between companion animal sentinel surveillance programs and human Lyme disease incidence. Together the results from these three studies close gaps in current knowledge surrounding accurate representation of surveillance data and provide the first account of ESDA techniques used to characterize occurrence of human Lyme disease in Kentucky, Ohio, and West Virginia.

Each of our studies were based upon investigation classes within spatial epidemiology with careful consideration of cartographic principles to avoid misrepresenting mapped information. Our first two investigations were an exercise in disease mapping and cluster analysis, while the third was consistent with an ecological analysis. It is important to note that while crude disease mapping and cluster analysis are useful in identifying patterns of disease intensity, they lack the ability to explain observed patterns (3). Therefore, approaches such as stratification and traditional confirmatory statistics were paired with ESDA techniques to provide further processing and complementary statistical information to avoid misrepresenting mapped information (4). Additionally, identical semiotic displays were employed utilizing the same

predefined classes and color schemes to represent LISA cluster map results. Scale for each of the studies was limited to the county level because of data availability. While some studies have criticized representation of surveillance information at scales as low as the county level (5), data on relevant potential confounders were collected no finer than the county-level.

Even with careful consideration of cartographic principles and pairing of statistical methodologies, it is important to note limitations to the project. Findings were based upon multiple exploratory spatial analyses of passively collected surveillance data. Therefore, results are potentially prone to misclassification bias stemming from inconsistent reporting influenced by varying presentation of patient signs and symptoms as well as unequal Lyme disease awareness (6-8). Additionally, the decision to designate a case as travel related was based upon self-reported travel history and most likely location for exposure. All of these are potentially prone to recall bias. Other important considerations are linked to volunteer biases potentially influencing the number of veterinary practices participating by county in the companion animal surveillance program (9). Unfortunately, we were unable to control for many of these information and selection biases due to the retrospective nature of our investigations.

Despite these limitations, ESDA tools and public health surveillance data offer a practical and efficient means to monitor occurrence and spread of disease (10). Overall, project findings:

- 1) Indicate a significant increasing time trend in the identification of high-high counties of Lyme disease in Maryland, Pennsylvania, Virginia, and West Virginia
- 2) Confirm travel associated case inclusion as a confounder of accurate representation of mapped Lyme disease surveillance data
- 3) Demonstrate the effectiveness of spatial regression models to evaluate the efficacy of animal sentinel programs in predicting occurrence of human Lyme disease.

These conclusions are consistent with previous research at the state and national level that observed Lyme disease cases clusters at state lines, and hypothesized this to result from disproportionate presence of infected ticks, and human behaviors such as travel (3, 11-16). Most notably, our findings support previous statements regarding the continually ongoing geographic expansion of high-risk Lyme disease designation (10). Future research is warranted to examine the extent of geographic influence other potential confounding variables have in accurate visualization of human Lyme disease incidence. Additionally, further research is critical to understanding the effect of increasing number of collection localities on spatial model efficiency for estimating associations between sentinel tick collection methodologies and human Lyme disease incidence.

LITERATURE CITED

1. Jacquez G. Spatial analysis in epidemiology: Nascent science or a failure of GIS? *J Geography Sys* 2000;2:91-7.
2. Killilea M, Swei A, Lane R, et al. Spatial Dynamics of Lyme Disease: A Review. *EcoHealth* 2008;5:167-95.
3. Szonyi B, Srinath I, Esteve-Gassent M, et al. Exploratory spatial analysis of Lyme disease in Texas -what can we learn from the reported cases? *BMC Public Health* 2015;15(924):1-8.
4. Lawson AB. *Statistical Methods in Spatial Epidemiology*. Chichester, NY: Wiley & Sons Ltd.; 2013.
5. Eisen L. ER. Need for Improved Methods to Collect and Present Spatial Epidemiologic Data for Vectorborne Diseases. *Emerging Infectious Diseases* 2007;13(12):1816-20.
6. Hayes E, Piesman J. How Can We Prevent Lyme Disease. *The New England Journal of Medicine* 2003;348:2424-30.
7. Bacon R, Kugeler K, Mead P. Surveillance for Lyme Disease --- United States, 1992--2006. *MMWR* 2008;57(SS10):1-9.
8. Brett M, Hinckley A, Zielinski-Gutierrez E, et al. U.S. healthcare providers' experience with Lyme and other tick-borne diseases. *Ticks and Tick-borne Diseases* 2014;5:404-8.
9. Johnson JL., Ginsberg HS., Zhioua E., et al. Passive Tick Surveillance, Dog Seropositivity, and Incidence of Human Lyme disease. *Vector-Borne and Zoonotic Diseases* 2004;4(2):137-42.
10. Kugeler KJ., Farley GM., Forrester JD., et al. Geographic Distribution and Expansion of Human Lyme Disease, United States. *Emerging Infectious Diseases* 2015;21(8):1455-57.
11. Bunnell J, Price S, Das A, et al. Geographic Information Systems and Spatial Analysis of Adult *Ixodes scapularis* (Acari: Ixodidae) in the Middle Atlantic Region of the U.S.A. *Journal of Medical Entomology* 2003;40(4):570-76.
12. Eisen R, Lane R, Fritz C, et al. Spatial Patterns of Lyme Disease Risk in California Based On Disease Incidence Data and Modeling of Vector-Tick Exposure. *American Journal of Tropical Medicine and Hygiene* 2006;75(4):669-76.
13. Glavanakov S, White D, Caraco T, et al. Lyme Disease in New York State: Spatial Pattern at a Regional Scale. *American Journal of Tropical Medicine and Hygiene* 2001;65(5):538-45.
14. Li J, Hong K, Duan Y, et al. Spatial and Temporal Emergence Pattern of Lyme Disease in Virginia. *American Journal Tropical Medicine and Hygiene* 2014;91(6):1166-72.
15. Brinkerhoff J, Gilliam W, Daines D. Lyme Disease, Virginia, USA, 2000-2011. *Emerging Infectious Diseases* 2015;20(10):1661-68.
16. Xue L, Scoglio C, McVey S, et al. Two Introductions of Lyme Disease into Connecticut: A Geospatial Analysis of Human Cases from 1984 to 2012. 2015:523-28.